

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number
WO 03/013226 A2

(51) International Patent Classification⁷: A01H

(21) International Application Number: PCT/US02/21837

(22) International Filing Date: 9 August 2002 (09.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/311,734 9 August 2001 (09.08.2001) US

(71) Applicant (for all designated States except US): CIBUS GENETICS [US/US]; 11180 Roselle Street, Suite 100, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): GOCAL, Greg

(74) Agents: SANDERS, James, M. et al.; Sughrue Mion, PLLC, 1010 El Camino Real, Menlo Park, CA 94025-4345 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS



WO 03/013226 A2

```

1  M A S S B L T K S I L O C T K P A
1 ATGGCGCTCT CCTCTCACTTC CAAATGATCT CTGGGATGCGA CCAAAACCGG
TACCGAGAA GAGGAGGAA CCTTAACTGAA GACCCATCAGT GGTGGGGGG
1- A B B S T P L P S E L R A L S S P A V
51 TCTTCTCTCT TTCTCTCTGT CGGGGGCTCC TCTCTCTCTCT TCTCCCGGCG
AACGAGAGAA MAAGAGGCGA CCTCTGGAGC ACCAGAGAGA AGAGGGCGGC
1- G L K R D Q D M I N H Q E I R P V
101 TTCAAGATTC TTCTCACTTC CAAACAGAGA AGAACTTCGCG CCAAGCTCTGG
AGCTCTTAA AGAAAGTGGT GTTGGGGCTT TTCTGAGAGC CGAACCAACCC
151 GGATGAGAA AGAGCTGATCT GATGCTTAAAT GGTCTGAGAA TTGGCTCTGG
CTTCAAGATTC TTCTCACTTC CAAACAGAGA AGAACTTCGCG CCAAGCTCTGG
1- V K V R A V A S T A E K A B I V L
201 GAGGGTAAAG CCTTCTCTGG CGGGGGCGGA GAAAGCTCTGG GAGGGTCTGG
CTTCAAGATTC CAAACAGAGA GGTCTGAGAA TTGGCTCTGG CCAAGCTCTGG
1- I Q P I R E S T G L I K P L I F P O S K
251 TTCAAGATTC TAAAGAGATTC TGGGGCTCTGG TAAAGCTCTGG TGGCTCTGG
AGCTGGGGTA ATCTCTCTGG ACCCCGGATC ATTTGGGGATC AGGAGGGTC
1- V V D H L L R B D D I R Y M L O A
301 TCTCTCTCTA ATGGGTTAACT GGTCTGGCTT GGTCTGAGAA AGGGAGGATC
AGGGAGGATC TAACTTAAAGA CAAAGAGAGA CCAAGCTCTGG TGGCTCTGG
1- A L K I L O L N V E T H E B E H R
351 TGATGTGAC AGCTTCTTACG AGCTGATCTA CAACTGATCTA ATGGCTGGTG
ACAGCTCTG TTGAGAGCT TCTGAGCTCTGG TAACTTAAAGA CAAAGAGAGA
401 CTTGGAGAT ATGGGCTCTG ATGGCTGGTG CTCACAGCTG AAGGGAGGATC
GCAAGCTCTG TAACTTAAAGA CAAAGAGAGA CCAAGCTCTGG TGGCTCTGG
1- A V V E G G G G V F P A S I D R K
451 GGTGTGGTGG AGGGAGGATG CGGGGGTAACT CAACTGATCTA TGATGTGAC
CAGAGAGAGA CTCAGCTCTGG TAACTTAAAGA CAAAGAGAGA CCAAGCTCTGG
1- K D O I E L Y L G H A Q Y A M R P I
501 GAGGGTGGATG AGCTTCTTACG TGGGGCTCTGG AGGGAGGATC ATGGCTGGTG
CTTCAAGATTC CTTGAGCTCTGG ACCGGGGATC TTCTGAGAGC CGAACCAACCC
1- L T A V T A A O G G R A S Y V L D
551 TTACCGGGCG AGTTAGCTG GCAAGCTCTGG ACCGGGGATC TTCTGAGAGC
AATGGGGCTCTGG CAACTGAGCTG GGTCTGAGAGC AGGGAGGATC
1- D V P R M R E R P I G D I V V O L
601 GGGGGCTCTGG GGTATGAGAGA GAAAGCTCTGG GGGGGATCTGG TTGGGGCTCTGG
CCCGGGAGGAG CCAAGCTCTGG TAACTTAAAGA CAAAGAGAGA CCAAGCTCTGG
1- I K L O G A D V E G T L O T I N C P E
651 TAACTGAGCTG GGTCTGAGAGC TTGGGGCTCTGG AGGGAGGATC ATGGCTGGTG
ATGGGGCTCTGG CAACTGAGCTG GGTCTGAGAGC AGGGAGGATC TTCTGAGAGC
1- F V R V R A N G G U L P G Q K V A L
701 CTGGGGCTCTGG CAACTGAGCTG GGTCTGAGAGC TTGGGGCTCTGG AGGGAGGATC
GCAAGCTCTGG TAACTTAAAGA CAAAGAGAGA CCAAGCTCTGG TGGCTCTGG
1- S G S I S O T L T A L L H A A P
751 TTCTGGAGCTCTGG TAACTTAAAGA CAAAGAGAGA CCAAGCTCTGG TGGCTCTGG
AGGGAGGATC ATGGGGCTCTGG CAACTGAGCTG GGTCTGAGAGC AGGGAGGATC

```

(57) Abstract: The present invention relates to the production of a non-transgenic plant resistant or tolerant to a herbicide of the phosphonomethylglycine family, e.g., glyphosate. The present invention also relates to the use of a recombinant oligonucleobase to make a desired mutation in the chromosomal or episomal sequences of a plant in the gene encoding for 5-enol pyruvylshikimate-3-phosphate synthase (EPSPS). The mutated protein, which substantially maintains the catalytic activity of the wild-type protein, allows for increased resistance or tolerance of the plant to a herbicide of the phosphonomethylglycine family, and allows for the substantially normal growth or development of the plant, its organs, tissues or cells as compared to the wild-type plant irrespective of the presence or absence of the herbicide. The present invention also relates to a non-transgenic plant cell in which the EPSPS gene has been mutated, a non-transgenic plant regenerated therefrom, as well as a plant resulting from a cross using a regenerated non-transgenic plant having a mutated EPSPS gene. The amino acids at the positions 126, 177, 207, 438, 479, 480 and/or 505 are changed to produce a mutant EPSPS gene product.

WO 03/013226 A2

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/013226

PCT/US02/21837

NON-TRANSGENIC HERBICIDE RESISTANT PLANTS**1. FIELD OF THE INVENTION**

The present invention relates to the production of a non-transgenic plant resistant or tolerant to a herbicide of the phosphonomethylglycine family, *e.g.*, glyphosate. The present invention also relates to the use of a recombinagenic oligonucleobase to make a desired mutation in the chromosomal or episomal sequences of a plant in the gene encoding for 5-enol pyruvylshikimate-3-phosphate synthase (EPSPS). The mutated protein, which substantially maintains the catalytic activity of the wild-type protein, allows for increased resistance or tolerance of the plant to a herbicide of the phosphonomethylglycine family, and allows for the substantially normal growth or development of the plant, its organs, tissues or cells as compared to the wild-type plant irrespective of the presence or absence of the herbicide. The present invention also relates to a non-transgenic plant cell in which the EPSPS gene has been mutated, a non-transgenic plant regenerated therefrom, as well as a plant resulting from a cross using a regenerated non-transgenic plant having a mutated EPSPS gene.

2. BACKGROUND TO THE INVENTION**2.1 PHOSPHONOMETHYLGlyCINE HERBICIDES**

Herbicide-tolerant plants may reduce the need for tillage to control weeds thereby effectively reducing soil erosion. One herbicide which is the subject of much investigation in this regard is N-phosphonomethylglycine, commonly referred to as glyphosate. Glyphosate inhibits the shikimic acid pathway which leads to the biosynthesis of aromatic compounds including amino acids, hormones and vitamins. Specifically, glyphosate curbs the conversion of phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid by inhibiting the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (hereinafter referred to as EPSP synthase or EPSPS). For purposes of the present invention, the term "glyphosate" includes any herbicidally effective form of N-phosphonomethylglycine (including any salt thereof), other forms which result in the production of the glyphosate anion in plants and any other herbicides of the phosphonomethylglycine family.

WO 03/013226

PCT/US02/21837

Tolerance of plants to glyphosate can be increased by introducing a mutant EPSPS gene having an alteration in the EPSPS amino acid coding sequence into the genome of the plant. Examples of some of the mutations in the EPSPS gene for inducing glyphosate tolerance are described in the following patents: U.S. Patent No. 5,310,667; U.S. Patent No. 5,866,775; U.S. Patent No. 5,312,910; U.S. Patent No. 5,145,783. These proposed mutations typically have a higher K_i for glyphosate than the wild-type EPSPS enzyme which confers the glyphosate-tolerant phenotype, but these variants are also characterized by a high K_m for PEP which makes the enzyme kinetically less efficient (Kishore et al., 1988, Ann. Rev. Biochem. 57:627-663; Schulz et al., 1984, Arch. Microbiol. 137: 121-123; Sost et al., 1984, FEBS Lett. 173: 238-241; Kishore et al., 1986, Fed. Proc. 45: 1506; Sost and Amrhein, 1990, Arch. Biochem. Biophys. 282: 433-436). Many mutations of the EPSPS gene are chosen so as to produce an EPSPS enzyme that is resistant to herbicides, but unfortunately, the EPSPS enzyme produced by the mutated EPSPS gene has a significantly lower enzymatic activity than the wild-type EPSPS. For example, the apparent K_m for PEP and the apparent K_i for glyphosate for the wild-type EPSPS from *E. coli* are 10 μ M and 0.5 μ M, while for a glyphosate-tolerant isolate having a single amino acid substitution of alanine for glycine at position 96, these values are 220 μ M and 4.0 mM, respectively. A number of glyphosate-tolerant EPSPS genes have been constructed by mutagenesis. Again, the glyphosate-tolerant EPSPS had lower catalytic efficiency (V_{max} / K_m), as shown by an increase in the K_m for PEP, and a slight reduction of the V_{max} of the wild-type plant enzyme (Kishore et al., 1988, Ann. Rev. Biochem. 57:627-663).

Since the kinetic constants of the variant enzymes are impaired with respect to PEP, it has been proposed that high levels of overproduction of the variant enzyme, 40-80 fold, would be required to maintain normal catalytic activity in plants in the presence of glyphosate (Kishore et al., 1988, Ann. Rev. Biochem. 57:627-663). It has been shown that glyphosate-tolerant plants can be produced by inserting into the genome of the plant the capacity to produce a higher level of EPSP synthase in the chloroplast of the cell (Shah et al., 1986, Science 233, 478-481), which enzyme is preferably glyphosate-tolerant (Kishore et al., 1988, Ann. Rev. Biochem. 57:627-663).

WO 03/013226

PCT/US02/21837

The introduction of the exogenous mutant EPSPS genes into plant cells is well documented. For example, according to U.S. Patent No. 4,545,060, to increase a plant's resistance to glyphosate, a gene coding for an EPSPS variant having at least one mutation that renders the enzyme more resistant to its competitive inhibitor, *i.e.*, glyphosate, is introduced into the plant genome. However, many complications and problems are associated with these examples. Many such mutations result in low expression of the mutated EPSPS gene product or result in an EPSPS gene product with significantly lower enzymatic activity as compared to the wild type. The low expression and/or low enzymatic activity of the mutated enzyme results in abnormally low levels of growth and development of the plant.

While such variants in the EPSP synthases have proved useful in obtaining transgenic plants tolerant to glyphosate, it would be increasingly beneficial to obtain a variant EPSPS gene product that is highly glyphosate-tolerant but still kinetically efficient, such that improved tolerance can be obtained with a wild-type expression level.

2.2 RECOMBINAGENIC OLIGONUCLEOBASES

Recombinagenic oligonucleobases and their use to effect genetic changes in eukaryotic cells are described in United States patent No. 5,565,350 to Kmiec (Kmiec I). Kmiec I teaches a method for introducing specific genetic alterations into a target gene. Kmiec I discloses, *inter alia*, recombinagenic oligonucleobases having two strands, in which a first strand contains two segments of at least 8 RNA-like nucleotides that are separated by a third segment of from 4 to about 50 DNA-like nucleotides, termed an "interposed DNA segment." The nucleotides of the first strand are base paired to DNA-like nucleotides of a second strand. The first and second strands are additionally linked by a segment of single stranded nucleotides so that the first and second strands are parts of a single oligonucleotide chain. Kmiec I further teaches a method for introducing specific genetic alterations into a target gene. According to Kmiec I, the sequences of the RNA segments are selected to be homologous, *i.e.*, identical, to the sequence of a first and a second fragment of the target gene. The sequence of the interposed DNA segment is homologous with the sequence of the target gene between the first and second fragment except for a region of difference, termed the "heterologous region." The heterologous region can effect an insertion or deletion, or can contain one or more bases that are mismatched with the sequence of target gene so as to effect

WO 03/013226

PCT/US02/21837

a substitution. According to Kmiec I, the sequence of the target gene is altered as directed by the heterologous region, such that the target gene becomes homologous with the sequence of the recombinagenic oligonucleobase. Kmiec I specifically teaches that ribose and 2'-O-methylribose, *i.e.*, 2'-methoxyribose, containing nucleotides can be used in recombinagenic oligonucleobases and that naturally-occurring deoxyribose-containing nucleotides can be used as DNA-like nucleotides.

U.S. Patent No. 5,731,181 to Kmiec (Kmiec II) specifically disclose the use of recombinagenic oligonucleobases to effect genetic changes in plant cells and discloses further examples of analogs and derivatives of RNA-like and DNA-like nucleotides that can be used to effect genetic changes in specific target genes. Other patents discussing the use of recombinagenic oligonucleobases include: U.S. Patent Nos. 5,756,325; 5,871,984; 5,760,012; 5,888,983; 5,795,972; 5,780,296; 5,945,339; 6,004,804; and 6,010,907 and in International Patent No. PCT/US00/23457; and in International Patent Publication Nos. WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; and WO 99/40789. Recombinagenic oligonucleobases include mixed duplex oligonucleotides, non-nucleotide containing molecules taught in Kmiec II and other molecules taught in the above-noted patents and patent publications.

Citation or identification of any reference in Section 2, or any section of this application shall not be construed as an admission that such reference is available as prior art to the present invention.

3. SUMMARY OF THE INVENTION

The present invention is directed to a non-transgenic plant or plant cell having one or more mutations in the EPSPS gene, which plant has increased resistance or tolerance to a member of the phosphonomethylglycine family and which plant exhibits substantially normal growth or development of the plant, its organs, tissues or cells, as compared to the corresponding wild-type plant or cell. The mutated gene produces a gene product having a substitution at one or more of the amino acid positions 126,177, 207, 438, 479,480 and 505 of the *Arabidopsis* EPSPS gene product or at an analogous amino acid position in an EPSPS homolog. The present invention is also directed to a non-transgenic plant having a mutation in the EPSPS gene, which plant is resistant to or has an increased tolerance to a member of

WO 03/013226

PCT/US02/21837

the phosphonomethylglycine family, *e.g.*, glyphosate, wherein the mutated EPSPS protein has substantially the same catalytic activity as compared to the wild-type EPSPS protein.

The present invention is also directed to a method for producing a non-transgenic plant having a mutated EPSPS gene that substantially maintains the catalytic activity of the wild-type protein irrespective of the presence or absence of a herbicide of the phosphonomethylglycine family. The method comprises introducing into a plant cell a recombinagenic oligonucleobase with a targeted mutation in the EPSPS gene and identifying a cell, seed, or plant having a mutated EPSPS gene.

Illustrative examples of a recombinagenic oligonucleobase are found in following patent publications, which are incorporated herein in their entirety by reference: U.S. Patent Nos. 5,565,350; 5,756,325; 5,871,984; 5,760,012; 5,731,181; 5,888,983; 5,795,972; 5,780,296; 5,945,339; 6,004,804; and 6,010,907 and in International Patent No. PCT/US00/23457; and in International Patent Publication Nos. WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; and WO 99/40789.

The plant can be of any species of dicotyledonous, monocotyledonous or gymnospermous plant, including any woody plant species that grows as a tree or shrub, any herbaceous species, or any species that produces edible fruits, seeds or vegetables, or any species that produces colorful or aromatic flowers. For example, the plant may be selected from a species of plant from the group consisting of canola, sunflower, tobacco, sugar beet, sweet potato, yam, cotton, maize, wheat, barley, rice, sorghum, tomato, mango, peach, apple, pear, strawberry, banana, melon, potato, carrot, lettuce, onion, soya spp, sugar cane, pea, peanut, field beans, poplar, grape, citrus, alfalfa, rye, oats, turf and forage grasses, flax, oilseed rape, cucumber, morning glory, balsam, pepper, eggplant, marigold, lotus, cabbage, daisy, carnation, tulip, iris, lily, and nut producing plants insofar as they are not already specifically mentioned.

The recombinagenic oligonucleobase can be introduced into a plant cell using any method commonly used in the art, including but not limited to, microcarriers (biolistic delivery), microfibers, electroporation, direct DNA uptake and microinjection.

The invention is also directed to the culture of cells mutated according to the methods of the present invention in order to obtain a plant that produces seeds, henceforth a

WO 03/013226

PCT/US02/21837

“fertile plant”, and the production of seeds and additional plants from such a fertile plant including descendant (progeny) plants that contain the mutated EPSPS gene.

The invention is further directed to a method of selectively controlling weeds in a field, the field comprising plants with the disclosed EPSPS gene alterations and weeds, the method comprising application to the field of a herbicide to which the said plants have been rendered resistant.

The invention is also directed to novel mutations in the EPSPS gene and resulting novel gene product that confer resistance or tolerance to a member of the phosphonomethylglycine family, *e.g.*, glyphosate, to a plant or wherein the mutated EPSPS has substantially the same enzymatic activity as compared to wild-type EPSPS.

3.1 DEFINITIONS

The invention is to be understood in accordance with the following definitions.

An oligonucleobase is a polymer of nucleobases, which polymer can hybridize by Watson-Crick base pairing to a DNA having the complementary sequence.

Nucleobases comprise a base, which is a purine, pyrimidine, or a derivative or analog thereof. Nucleobases include peptide nucleobases, the subunits of peptide nucleic acids, and morpholine nucleobases as well as nucleosides and nucleotides. Nucleosides are nucleobases that contain a pentosefuranosyl moiety, *e.g.*, an optionally substituted riboside or 2'-deoxyriboside. Nucleosides can be linked by one of several linkage moieties, which may or may not contain a phosphorus. Nucleosides that are linked by unsubstituted phosphodiester linkages are termed nucleotides.

An oligonucleobase chain has a single 5' and 3' terminus, which are the ultimate nucleobases of the polymer. A particular oligonucleobase chain can contain nucleobases of all types. An oligonucleobase compound is a compound comprising one or more oligonucleobase chains that are complementary and hybridized by Watson-Crick base pairing. Nucleobases are either deoxyribo-type or ribo-type. Ribo-type nucleobases are pentosefuranosyl containing nucleobases wherein the 2' carbon is a methylene substituted with a hydroxyl, alkyloxy or halogen. Deoxyribo-type nucleobases are nucleobases other than ribo-type nucleobases and include all nucleobases that do not contain a pentosefuranosyl moiety.

WO 03/013226

PCT/US02/21837

An oligonucleobase strand generically includes both oligonucleobase chains and segments or regions of oligonucleobase chains. An oligonucleobase strand has a 3' end and a 5' end. When a oligonucleobase strand is coextensive with a chain, the 3' and 5' ends of the strand are also 3' and 5' termini of the chain.

According to the present invention, substantially normal growth of a plant, plant organ, plant tissue or plant cell is defined as a growth rate or rate of cell division of the plant, plant organ, plant tissue, or plant cell that is at least 35%, at least 50%, at least 60%, or at least 75% of the growth rate or rate of cell division in a corresponding plant, plant organ, plant tissue or plant cell expressing the wild type EPSPS protein.

According to the present invention, substantially normal development of a plant, plant organ, plant tissue or plant cell is defined as the occurrence of one or more developmental events in the plant, plant organ, plant tissue or plant cell that are substantially the same as those occurring in a corresponding plant, plant organ, plant tissue or plant cell expressing the wild type EPSPS protein.

According to the present invention plant organs include, but are not limited to, leaves, stems, roots, vegetative buds, floral buds, meristems, embryos, cotyledons, endosperm, sepals, petals, pistils, carpels, stamens, anthers, microspores, pollen, pollen tubes, ovules, ovaries and fruits, or sections, slices or discs taken therefrom. Plant tissues include, but are not limited to, callus tissues, ground tissues, vascular tissues, storage tissues, meristematic tissues, leaf tissues, shoot tissues, root tissues, gall tissues, plant tumor tissues, and reproductive tissues. Plant cells include, but are not limited to, isolated cells with cell walls, variously sized aggregates thereof, and protoplasts.

Plants are substantially "tolerant" to glyphosate when they are subjected to it and provide a dose/response curve which is shifted to the right when compared with that provided by similarly subjected non-tolerant like plant. Such dose/response curves have "dose" plotted on the X-axis and "percentage kill", "herbicidal effect", etc., plotted on the y-axis. Tolerant plants will require more herbicide than non-tolerant like plants in order to produce a given herbicidal effect. Plants which are substantially "resistant" to the glyphosate exhibit few, if any, necrotic, lytic, chlorotic or other lesions, when subjected to glyphosate at concentrations and rates which are typically employed by the agrochemical community to kill

WO 03/013226

PCT/US02/21837

weeds in the field. Plants which are resistant to a herbicide are also tolerant of the herbicide. The terms "resistant" and "tolerant" are to be construed as "tolerant and/or resistant" within the context of the present application.

The term "EPSPS homolog" or any variation therefore refers to an EPSPS gene or EPSPS gene product found in another plant species that performs the same or substantially the same biological function as the EPSPS genes disclosed herein and where the nucleic acid sequences or polypeptide sequences (of the EPSPS gene product) are said to be "identical" or at least 50 % similar (also referred to as 'percent identity' or' substantially identical') as described below. Two polynucleotides or polypeptides are identical if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence as described below. The terms "identical" or "percent identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. For polypeptides where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a 'score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated according to, e.g., the algorithm of Meyers & Miller, Computer Applic. Biol. Sci. 4: 1 1-17 (1988) e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA).

The phrases "substantially identical," and "percent identity" in the context of two nucleic acids or polypeptides, refer to sequences or subsequences that have at least 50%, advantageously 60%, preferably 70%, more preferably 80%, and most preferably 90-95% nucleotide or amino acid residue identity when aligned for maximum correspondence over a

WO 03/013226

PCT/US02/21837

comparison window as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence, which has substantial sequence or subsequence complementarity when the test sequence has substantial identity to a reference sequence.

One of skill in the art will recognize that two polypeptides can also be "substantially identical" if the two polypeptides are immunologically similar. Thus, overall protein structure may be similar while the primary structure of the two polypeptides display significant variation. Therefore a method to measure whether two polypeptides are substantially identical involves measuring the binding of monoclonal or polyclonal antibodies to each polypeptide. Two polypeptides are substantially identical if the antibodies specific for a first polypeptide bind to a second polypeptide with an affinity of at least one third of the affinity for the first polypeptide. For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), by software for alignments such as VECTOR NTI Version #6 by InforMax, Inc. MD, USA, by the procedures described in ClustalW, Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position – specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22:4673-4680 or by visual inspection (see generally, *Protocols in Molecular Biology*, F.M. Ausubel et al., eds., Current Protocols, a

WO 03/013226

PCT/US02/21837

joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)).

Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1990) *J. Mol. Biol.* 215: 403-410 and Altschul et al. (1977) *Nucleic Acids Res.* 25 : 33 89-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a word length (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a word length (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff& Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)). In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an

WO 03/013226

PCT/US02/21837

indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

4. BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 is the cDNA sequence and the amino acid sequence of *Arabidopsis thaliana* EPSPS gene. The underlined nucleotide and amino acid residues are the targeted residues. (GenBank accession number AY040065)

FIG. 2 shows (1) a table of the present EPSPS mutants by comparing the mutated amino acid positions in the *E. coli* AroA gene product with the *Arabidopsis* mutations and (2) a list of (a-i) the *Arabidopsis thaliana* wild-type and mutant EPSPS nucleotide sequences in the region of the mutations where the upper sequence represents the wild-type sequence and the lower sequence represents the mutated sequence. The lower case nucleotides represent the mutation.

FIG. 3 is an alignment of the amino acid sequences of various EPSPS gene products performed by VECTOR NTI. The sequences were aligned using the CLUSTAL W methodology. Residues in an alignment are colored according to the following scheme:

black on window default color -- non-similar residues;

blue on cyan -- consensus residue derived from a block of similar residues at a given position;

black on green -- consensus residue derived from the occurrence of greater than 50% of a single residue at a given position;

red on yellow -- consensus residue derived from a completely conserved residue at a given position;

green on window default color -- residue weakly similar to consensus residue at given position.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to a non-transgenic plant or plant cell having a mutation in the EPSPS gene, which plant has increased resistance or tolerance to a member

WO 03/013226

PCT/US02/21837

of the phosphonomethylglycine family and which plant exhibits substantially normal growth or development of the plant, its organs, tissues or cells, as compared to the corresponding wild-type plant or cell. The present invention is also directed to a non-transgenic plant having a mutation in the EPSPS gene, which plant is resistant to or has an increased tolerance to a member of the phosphonomethylglycine family, *e.g.*, glyphosate, wherein the mutated EPSPS protein has substantially the same catalytic activity as compared to the wild-type EPSPS protein.

The present invention is also directed to a method for producing a non-transgenic plant having a mutated EPSPS gene that substantially maintains the catalytic activity of the wild-type protein irrespective of the presence or absence of a herbicide of the phosphonomethylglycine family. The method comprises introducing into a plant cell a recombinagenic oligonucleobase with a targeted mutation in the EPSPS gene and identifying a cell, seed, or plant having a mutated EPSPS gene.

Illustrative examples of a recombinagenic oligonucleobase is found in following patent publications, which are incorporated in their entirety by reference herein: U.S. Patent Nos. 5,565,350; 5,756,325; 5,871,984; 5,760,012; 5,731,181; 5,888,983; 5,795,972; 5,780,296; 5,945,339; 6,004,804; and 6,010,907 and in International Patent No. PCT/US00/23457; and in International Patent Publication Nos. WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; and WO 99/40789.

The plant can be of any species of dicotyledonous, monocotyledonous or gymnospermous plant, including any woody plant species that grows as a tree or shrub, any herbaceous species, or any species that produces edible fruits, seeds or vegetables, or any species that produces colorful or aromatic flowers. For example, the plant may be selected from a species of plant from the group consisting of canola, sunflower, tobacco, sugar beet, cotton, maize, wheat, barley, rice, sorghum, tomato, mango, peach, apple, pear, strawberry, banana, melon, potato, sweet potato, yam, carrot, lettuce, onion, soya spp, sugar cane, pea, peanut, field beans, poplar, grape, citrus, alfalfa, rye, oats, lentils, turf and forage grasses, eucalyptus, flax, oilseed rape, cucumber, morning glory, balsam, pepper, eggplant, marigold, lotus, cabbage, daisy, carnation, tulip, iris, lily, and nut producing plants insofar as they are not already specifically mentioned.

WO 03/013226

PCT/US02/21837

The recombinagenic oligonucleobase can be introduced into a plant cell using any method commonly used in the art, including but not limited to, microcarriers (biolistic delivery), microfibers, electroporation, direct DNA uptake (including polyethylene mediated DNA uptake) and microinjection.

The invention is also directed to the culture of cells mutated according to the methods of the present invention in order to obtain a plant that produces seeds, henceforth a "fertile plant", and the production of seeds and additional plants from such a fertile plant including descendant (progeny) plants that contain the mutated EPSPS gene.

The invention is further directed to a method of selectively controlling weeds in a field, the field comprising plants with the disclosed EPSPS gene alterations and weeds, the method comprising application to the field of a phosphonomethylglycine herbicide to which the said plants have been rendered resistant.

The invention is also directed to novel mutations in the EPSPS gene and gene product that confer resistance or tolerance to a member of the phosphonomethylglycine family, *e.g.*, glyphosate, to a plant or wherein the mutated EPSPS has substantially the same enzymatic activity as compared to wild-type EPSPS.

5.1 RECOMBINAGENIC OLIGONUCLEOBASES

The invention can be practiced with recombinagenic oligonucleobases having the conformations and chemistries described in United States patent No. 5,565,350 to Kmiec (Kmiec I) and U.S. patent No. 5,731,181 (Kmiec II) gene, which are incorporated herein by reference. Kmiec I teaches a method for introducing specific genetic alterations into a target gene. The recombinagenic oligonucleobases in Kmiec I and/or Kmiec II contain two complementary strands, one of which contains at least one segment of RNA-type nucleotides (an "RNA segment") that are base paired to DNA-type nucleotides of the other strand.

Kmiec II discloses that purine and pyrimidine base-containing non-nucleotides can be substituted for nucleotides. U.S. Patent Nos. 5,756,325; 5,871,984; 5,760,012; 5,888,983; 5,795,972; 5,780,296; 5,945,339; 6,004,804; and 6,010,907 and in International Patent No. PCT/US00/23457; and in International Patent Publication Nos. WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; and WO 99/40789, which are each hereby incorporated in their entirety, disclose additional recombinagenic molecules that can be used

WO 03/013226

PCT/US02/21837

for the present invention. The term "recombinagenic oligonucleobase" is used herein to denote the molecules that can be used in the methods of the present invention and include mixed duplex oligonucleotides, non-nucleotide containing molecules taught in Kmiec II, single stranded oligodeoxynucleotides and other recombinagenic molecules taught in the above noted patents and patent publications.

In one embodiment, the recombinagenic oligonucleobase is a mixed duplex oligonucleotide in which the RNA-type nucleotides of the mixed duplex oligonucleotide are made RNase resistant by replacing the 2'-hydroxyl with a fluoro, chloro or bromo functionality or by placing a substituent on the 2'-O. Suitable substituents include the substituents taught by the Kmiec II. Alternative substituents include the substituents taught by U.S. Patent No. 5,334,711 (Sproat) and the substituents taught by patent publications EP 629 387 and EP 679 657 (collectively, the Martin Applications), which are incorporated herein by reference. As used herein, a 2' -fluoro, chloro or bromo derivative of a ribonucleotide or a ribonucleotide having a 2'-OH substituted with a substituent described in the Martin Applications or Sproat is termed a "2'-Substituted Ribonucleotide." As used herein the term "RNA-type nucleotide" means a 2'-hydroxyl or 2'-Substituted Nucleotide that is linked to other nucleotides of a mixed duplex oligonucleotide by an unsubstituted phosphodiester linkage or any of the non-natural linkages taught by Kmiec I or Kmiec II. As used herein the term "deoxyribo-type nucleotide" means a nucleotide having a 2'-H, which can be linked to other nucleotides of a MDON by an unsubstituted phosphodiester linkage or any of the non-natural linkages taught by Kmiec I or Kmiec II.

In a particular embodiment of the present invention, the recombinagenic oligonucleobase is a mixed duplex oligonucleotide that is linked solely by unsubstituted phosphodiester bonds. In alternative embodiments, the linkage is by substituted phosphodiesters, phosphodiester derivatives and non-phosphorus-based linkages as taught by Kmiec II. In yet another embodiment, each RNA-type nucleotide in the mixed duplex oligonucleotide is a 2'-Substituted Nucleotide. Particularly preferred embodiments of 2'-Substituted Ribonucleotides are 2'-fluoro, 2'-methoxy, 2'-propyloxy, 2'-allyloxy, 2'-hydroxylethoxy, 2'-methoxyethoxy, 2'-fluoropropoxy and 2'-trifluoropropoxy substituted ribonucleotides. More preferred embodiments of 2'-Substituted Ribonucleotides

WO 03/013226

PCT/US02/21837

are 2'-fluoro, 2'-methoxy, 2'-methoxyethoxy, and 2'-allyloxy substituted nucleotides. In another embodiment the mixed duplex oligonucleotide is linked by unsubstituted phosphodiester bonds.

Although mixed duplex oligonucleotide having only a single type of 2'-substituted RNA-type nucleotide are more conveniently synthesized, the methods of the invention can be practiced with mixed duplex oligonucleotides having two or more types of RNA-type nucleotides. The function of an RNA segment may not be affected by an interruption caused by the introduction of a deoxynucleotide between two RNA-type trinucleotides, accordingly, the term RNA segment encompasses such an "interrupted RNA segment." An uninterrupted RNA segment is termed a contiguous RNA segment. In an alternative embodiment an RNA segment can contain alternating RNase-resistant and unsubstituted 2'-OH nucleotides. The mixed duplex oligonucleotides preferably have fewer than 100 nucleotides and more preferably fewer than 85 nucleotides, but more than 50 nucleotides. The first and second strands are Watson-Crick base paired. In one embodiment the strands of the mixed duplex oligonucleotide are covalently bonded by a linker, such as a single stranded hexa, penta or tetranucleotide so that the first and second strands are segments of a single oligonucleotide chain having a single 3' and a single 5' end. The 3' and 5' ends can be protected by the addition of a "hairpin cap" whereby the 3' and 5' terminal nucleotides are Watson-Crick paired to adjacent nucleotides. A second hairpin cap can, additionally, be placed at the junction between the first and second strands distant from the 3' and 5' ends, so that the Watson-Crick pairing between the first and second strands is stabilized.

The first and second strands contain two regions that are homologous with two fragments of the target EPSPS gene, *i.e.*, have the same sequence as the target gene. A homologous region contains the nucleotides of an RNA segment and may contain one or more DNA-type nucleotides of connecting DNA segment and may also contain DNA-type nucleotides that are not within the intervening DNA segment. The two regions of homology are separated by, and each is adjacent to, a region having a sequence that differs from the sequence of the target gene, termed a "heterologous region." The heterologous region can contain one, two or three mismatched nucleotides. The mismatched nucleotides can be contiguous or alternatively can be separated by one or two nucleotides that are homologous

WO 03/013226

PCT/US02/21837

with the target gene. Alternatively, the heterologous region can also contain an insertion or one, two, three or of five or fewer nucleotides. Alternatively, the sequence of the mixed duplex oligonucleotide may differ from the sequence of the target gene only by the deletion of one, two, three, or five or fewer nucleotides from the mixed duplex oligonucleotide. The length and position of the heterologous region is, in this case, deemed to be the length of the deletion, even though no nucleotides of the mixed duplex oligonucleotide are within the heterologous region. The distance between the fragments of the target gene that are complementary to the two homologous regions is identically the length of the heterologous region when a substitution or substitutions is intended. When the heterologous region contains an insertion, the homologous regions are thereby separated in the mixed duplex oligonucleotide farther than their complementary homologous fragments are in the gene, and the converse is applicable when the heterologous region encodes a deletion.

The RNA segments of the mixed duplex oligonucleotides are each a part of a homologous region, *i.e.*, a region that is identical in sequence to a fragment of the target gene, which segments together preferably contain at least 13 RNA-type nucleotides and preferably from 16 to 25 RNA-type nucleotides or yet more preferably 18-22 RNA-type nucleotides or most preferably 20 nucleotides. In one embodiment, RNA segments of the homology regions are separated by and adjacent to, *i.e.*, “connected by” an intervening DNA segment. In one embodiment, each nucleotide of the heterologous region is a nucleotide of the intervening DNA segment. An intervening DNA segment that contains the heterologous region of a mixed duplex oligonucleotide is termed a “mutator segment.”

The change to be introduced into the target EPSPS gene is encoded by the heterologous region. The change to be introduced into the EPSPS gene may be a change in one or more bases of the EPSPS gene sequence or the addition or deletion of one or more bases.

In another embodiment of the present invention, the recombinagenic oligonucleobase is a single stranded oligodeoxynucleotide mutational vector or SSOMV, which is disclosed in International Patent Application PCT/US00/23457, which is incorporated herein by reference in its entirety. The sequence of the SSOMV is based on the same principles as the mutational vectors described in U.S. Patent Nos. 5,756,325; 5,871,984;

WO 03/013226

PCT/US02/21837

5,760,012; 5,888,983; 5,795,972; 5,780,296; 5,945,339; 6,004,804; and 6,010,907 and in International Publication Nos. WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; and WO 99/40789. The sequence of the SSOMV contains two regions that are homologous with the target sequence separated by a region that contains the desired genetic alteration termed the mutator region. The mutator region can have a sequence that is the same length as the sequence that separates the homologous regions in the target sequence, but having a different sequence. Such a mutator region can cause a substitution. Alternatively, the homologous regions in the SSOMV can be contiguous to each other, while the regions in the target gene having the same sequence are separated by one, two or more nucleotides. Such a SSOMV causes a deletion from the target gene of the nucleotides that are absent from the SSOMV. Lastly, the sequence of the target gene that is identical to the homologous regions may be adjacent in the target gene but separated by one two or more nucleotides in the sequence of the SSOMV. Such an SSOMV causes an insertion in the sequence of target gene.

The nucleotides of the SSOMV are deoxyribonucleotides that are linked by unmodified phosphodiester bonds except that the 3' terminal and/or 5' terminal internucleotide linkage or alternatively the two 3' terminal and/or 5' terminal internucleotide linkages can be a phosphorothioate or phosphoamidate. As used herein an internucleotide linkage is the linkage between nucleotides of the SSOMV and does not include the linkage between the 3' end nucleotide or 5' end nucleotide and a blocking substituent, see *supra*. In a specific embodiment the length of the SSOMV is between 21 and 55 deoxynucleotides and the lengths of the homology regions are, accordingly, a total length of at least 20 deoxynucleotides and at least two homology regions should each have lengths of at least 8 deoxynucleotides.

The SSOMV can be designed to be complementary to either the coding or the non-coding strand of the target gene. When the desired mutation is a substitution of a single base, it is preferred that both the mutator nucleotide be a pyrimidine. To the extent that is consistent with achieving the desired functional result it is preferred that both the mutator nucleotide and the targeted nucleotide in the complementary strand be pyrimidines.

WO 03/013226

PCT/US02/21837

Particularly preferred are SSOMV that encode transversion mutations, *i.e.*, a C or T mutator nucleotide is mismatched, respectively, with a C or T nucleotide in the complementary strand.

In addition to the oligodeoxynucleotide the SSOMV can contain a 5' blocking substituent that is attached to the 5' terminal carbons through a linker. The chemistry of the linker is not critical other than its length, which should preferably be at least 6 atoms long and that the linker should be flexible. A variety of non-toxic substituents such as biotin, cholesterol or other steroids or a non-intercalating cationic fluorescent dye can be used.

Particularly preferred as reagents to make SSOMV are the reagents sold as Cy3™ and Cy5™ by Glen Research, Sterling VA, which are blocked phosphoroamidites that upon incorporation into an oligonucleotide yield 3,3,3',3'-tetramethyl N,N'-isopropyl substituted indomonocarbocyanine and indodicarbocyanine dyes, respectively. Cy3 is the most preferred. When the indocarbocyanine is N-oxyalkyl substituted it can be conveniently linked to the 5' terminal of the oligodeoxynucleotide through as a phosphodiester with a 5' terminal phosphate. The chemistry of the dye linker between the dye and the oligodeoxynucleotide is not critical and is chosen for synthetic convenience. When the commercially available Cy3 phosphoramidite is used as directed the resulting 5' modification consists of a blocking substituent and linker together which are a N-hydroxypropyl, N'-phosphatidylpropyl 3,3,3',3'-tetramethyl indomonocarbocyanine.

In the preferred embodiment the indocarbocyanine dye is tetra substituted at the 3 and 3' positions of the indole rings. Without limitation as to theory these substitutions prevent the dye from being an intercalating dye. The identity of the substituents at these positions are not critical. The SSOMV can in addition have a 3' blocking substituent. Again the chemistry of the 3' blocking substituent is not critical.

5.2 THE LOCATION AND TYPE OF MUTATION

INTRODUCED INTO THE EPSPS GENE

In one embodiment of the present invention, the *Arabidopsis thaliana* EPSPS gene and corresponding EPSPS gene product (enzyme) (see Figure 1) comprises a mutation at one or more amino acid residues selected from the group consisting of D₁₂₆, R₂₀₇, R₄₃₈, H₄₇₉, R₄₈₀, G₁₇₇ and K₅₀₅ or at an analogous position in an EPSPS homolog, and the mutation

WO 03/013226

PCT/US02/21837

results in one or more of the following amino acid substitutions in the EPSPS enzyme in comparison with the wild-type sequence:

- (i) Asp₁₂₆ - Glu
- (ii) Arg₂₀₇ - Glu
- (iii) Arg₄₃₈ - Lys
- (iv) His₄₇₉ - Arg or Leu
- (v) His₄₇₉R₄₈₀ - Arg₄₇₉Lys₄₈₀
- (vi) Gly₁₇₇ - Met or Ser
- (vii) Lys₅₀₅ - Arg

Alternatively, and/or additionally, the mutation may result in the replacement of any amino acid at positions corresponding to 126, 177, 207, 438, 479, 480 (if amino acid 479 is replaced) and 505 with respect to the EPSPS protein depicted in Figure 1.

In specific embodiments of the present invention, the EPSPS gene is mutated at amino acid position 126 in which Asp is replaced by Glu. Another specific embodiment is the substitution of Arg at amino acid position 207 by Glu. A further specific embodiment comprises a mutation at amino acid position 480 in which Arg is replaced by Lys, plus the additional substitution of His at amino acid position 479 by Arg. Other specific embodiments of the present invention are directed to mutations at amino acid position 438, in which Arg is replaced by Lys; amino acid position 479, in which His is replaced by Arg or Leu; amino acid position 177 in which Gly is substituted by Ser or Met; and amino acid position 505 in which Lys is replaced by Arg.

The foregoing mutations in the EPSPS gene are seen in the *Arabidopsis thaliana* EPSPS gene and protein sequences in FIG. 1. The present invention also encompasses mutant EPSPS genes of other plant species (homologs). However, due to variations in the EPSPS genes of different species, the position number of the amino acid residue to be changed in one species may be different in another species. Nevertheless, the analogous position is readily identified by one of skill in the art by sequence homology. For example, Figure 3 shows the aligned amino acid sequences of homologs of the EPSPS gene in various organisms including, *Arabidopsis thaliana*, *Zea mays*, *Petunia hybrida*, *N. tabacum*, tomato and *Brassica napus*. Thus, the analogous positions in *Zea mays* are Asp₅₁,

WO 03/013226

PCT/US02/21837

Gly₁₀₁, Arg₁₃₁, Arg₃₆₂, His₄₀₃, Arg₄₀₄ and Lys₄₂₉. Thus, the *Zea mays* EPSPS amino acid sequence is mutated at one or more of the following amino acid positions and results in one or more of the following substitutions:

- (i) Asp₅₁ - Glu
- (ii) Gly₁₀₁ - Ser or Met
- (iii) Arg₁₃₁ - Glu
- (iv) Arg₃₆₂ - Lys
- (v) His₄₀₃ - Leu or Arg
- (vi) His₄₀₃Arg₄₀₄ - Arg₄₀₃Lys₄₀₄
- (vii) Lys₄₂₉ - Arg

In *Brassica napus*, the analogous amino acid positions are D₁₂₂, R₂₀₃, R₄₃₄, H₄₇₅, R₄₇₆, G₁₇₃ and K₅₀₁. Thus, the *Brassica napus* EPSPS amino acid sequence is mutated at one or more of the following amino acid positions and results in one or more of the following substitutions:

- (i) Asp₁₂₂ - Glu
- (ii) Arg₂₀₃ - Glu
- (iii) Arg₄₃₄ - Lys
- (iv) His₄₇₅ - Leu or Arg
- (v) His₄₇₅Arg₄₇₆ - Arg₄₇₅Lys₄₇₆
- (vi) Gly₁₇₃ - Met or Ser
- (vii) Lys₅₀₁ - Arg

In *Petunia hybrida* the analogous positions are D₁₂₂, R₂₀₃, R₄₃₄, H₄₇₅, R₄₇₆, G₁₇₃ and K₅₀₁. Thus, the *Petunia hybrida* EPSPS amino acid sequence is mutated at one or more of the following amino acid positions and results in one or more of the following substitutions:

- (i) Asp₁₂₂ - Glu
- (ii) Arg₂₀₃ - Glu
- (iii) Arg₄₃₄ - Lys
- (iv) His₄₇₅ - Leu or Arg
- (v) His₄₇₅Arg₄₇₆ - Arg₄₇₅Lys₄₇₆

(vi) Gly₁₇₃ - Met or Ser

(vii) Lys₅₀₁ - Arg

5.3 THE DELIVERY OF RECOMBINAGENIC

OLIGONUCLEOBASES INTO PLANT CELLS

Any commonly known method can be used in the methods of the present invention to transform a plant cell with a recombinagenic oligonucleobases. Illustrative methods are listed below.

5.3.1 MICROCARRIERS AND MICROFIBERS

The use of metallic microcarriers (microspheres) for introducing large fragments of DNA into plant cells having cellulose cell walls by projectile penetration is well known to those skilled in the relevant art (henceforth biolistic delivery). United States Patent Nos. 4,945,050; 5,100,792 and 5,204,253 describe general techniques for selecting microcarriers and devices for projecting them. US Patents 5,484,956 and 5,489,520 describe the preparation of fertile transgenic corn using microprojectile bombardment of corn callus tissue. The biolistic techniques are also used in transforming immature corn embryos.

Specific conditions for using microcarriers in the methods of the present invention are described in International Publication WO 99/07865. In an illustrative technique, ice cold microcarriers (60 mg/ml), mixed duplex oligonucleotide (60 mg/ml) 2.5 M CaCl₂ and 0.1 M spermidine are added in that order; the mixture is gently agitated, e.g., by vortexing, for 10 minutes and let stand at room temperature for 10 minutes, whereupon the microcarriers are diluted in 5 volumes of ethanol, centrifuged and resuspended in 100% ethanol. Good results can be obtained with a concentration in the adhering solution of 8-10 µg/µl microcarriers, 14-17 µg/ml mixed duplex oligonucleotide, 1.1-1.4 M CaCl₂ and 18-22 mM spermidine. Optimal results were observed under the conditions of 8 µg/µl microcarriers, 16.5 µg/ml mixed duplex oligonucleotide, 1.3 M CaCl₂ and 21 mM spermidine.

Recombinagenic oligonucleobases can also be introduced into plant cells for the practice of the present invention using microfibers to penetrate the cell wall and cell membrane. U.S. Patent No. 5,302,523 to Coffee et al. describes the use of 30 x 0.5 µm and 10 x 0.3 µm silicon carbide fibers to facilitate transformation of suspension maize cultures of

WO 03/013226

PCT/US02/21837

Black Mexican Sweet. Any mechanical technique that can be used to introduce DNA for transformation of a plant cell using microfibers can be used to deliver recombinagenic oligonucleobases for use in making the present EPSPS mutants. The process disclosed by Coffee et al in US 5,302,523 can be employed with regenerable plant cell materials to introduce the present recombinagenic oligonucleobases to effect the mutation of the EPSPS gene whereby a whole mutated plant can be recovered that exhibits the glyphosate resistant phenotype.

An illustrative technique for microfiber delivery of a recombinagenic oligonucleobase is as follows: Sterile microfibers (2 µg) are suspended in 150 µl of plant culture medium containing about 10 µg of a mixed duplex oligonucleotide. A suspension culture is allowed to settle and equal volumes of packed cells and the sterile fiber/nucleotide suspension are vortexed for 10 minutes and plated. Selective media are applied immediately or with a delay of up to about 120 hours as is appropriate for the particular trait.

5.3.2 ELECTROPORATION

In an alternative embodiment, the recombinagenic oligonucleobases can be delivered to the plant cell by electroporation of a protoplast derived from a plant part. The protoplasts are formed by enzymatic treatment of a plant part, such as a leaf, according to techniques well known to those skilled in the art. *See, e.g., Gallois et al., 1996, in Methods in Molecular Biology 55:89-107, Humana Press, Totowa, NJ; Kipp et al., 1999, in Methods in Molecular Biology 133:213-221, Humana Press, Totowa, NJ.* The protoplasts need not be cultured in growth media prior to electroporation. Illustrative conditions for electroporation are 3×10^5 protoplasts in a total volume of 0.3 ml with a concentration of recombinagenic oligonucleobase of between 0.6 - 4 µg/mL.

Recombinagenic oligonucleobases can also be introduced into microspores by electroporation. Upon release of the tetrad, the microspore is uninucleate and thin-walled. It begins to enlarge and develops a germ pore before the exine forms. A microspore at this stage is potentially more amenable to transformation with exogenous DNA than other plant cells. In addition, microspore development can be altered *in vitro* to produce either haploid embryos or embryogenic callus that can be regenerated into plants (Coumaris et al., *Plant Cell Rep.* 7:618-621, 1989; Datta et al., *Plant Sci.* 67:83-88, 1990; Maheshwari et al., *Am. J. Bot.*

WO 03/013226

PCT/US02/21837

69:865-879, 1982; Schaeffer, *Adv. In Cell Culture* 7:161-182, 1989; Swanson et al., *Plant Cell Rep.* 6:94-97, 1987). Thus, transformed microspores can be regenerated directly into haploid plants or dihaploid fertile plants upon chromosome doubling by standard methods. See also co-pending application United States Serial Number 09/680,858 entitled Compositions and Methods for Plant Genetic Modification which is incorporated herein by reference.

Microspore electroporation can be practiced with any plant species for which microspore culture is possible, including but not limited to plants in the families Gramineae, Leguminosae, Cruciferaceae, Solanaceae, Cucurbitaceae, Rosaceae, Poaceae, Lilaceae, Rutaceae, Vitaceae, including such species as corn (*Zea mays*), wheat (*Triticum aestivum*), rice (*Oryza sativa*), oats, barley, canola (*Brassica napus*, *Brassica rapa*, *Brassica oleracea*, and *Brassica juncea*), cotton (*Gossypium hirsutum* L.), various legume species (e.g., soybean [*Glycine max*], pea [*Pisum sativum*], etc.), grapes [*Vitis vinifera*], and a host of other important crop plants. Microspore embryogenesis, both from anther and microspore culture, has been described in more than 170 species, belonging to 68 genera and 28 families of dicotyledons and monocotyledons (Raghavan, *Embryogenesis in Agniosperms: A Developmental and Experimental Study*, Cambridge University Press, Cambridge, England, 1986; Raghavan, *Cell Differentiation* 21:213-226, 1987; Raemakers et al., *Euphytica* 81:93-107, 1995). For a detailed discussion of microspore isolation, culture, and regeneration of double haploid plants from microspore-derived embryos [MDE] in *Brassica napus* L., see Nehlin, *The Use of Rapeseed (Brassica napus L.) Microspores as a Tool for Biotechnological Applications*, doctoral thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 1999; also Nehlin et al., *Plant Sci.* 111:219-227, 1995, and Nehlin et al., *Plant Sci.* 111:219-227, 1995). Chromosome doubling from microspore or anther culture is a well-established technique for production of double-haploid homozogous plant lines in several crops (Heberle-Bors et al., *In vitro pollen cultures: Progress and perspectives*. In: *Pollen Biotechnology. Gene expression and allergen characterization*, vol. 85-109, ed. Mohapatra, S. S., and Knox, R. B., Chapman and Hall, New York, 1996).

Microspore electroporation methods are described in Jardinaud et al., *Plant Sci.* 93:177-184, 1993, and Fennell and Hauptman, *Plant Cell Reports* 11:567-570, 1992.

WO 03/013226

PCT/US02/21837

Methods for electroporation of MDON into plant protoplasts can also be adapted for use in microspore electroporation.

5.3.3 WHISKERS AND MICROINJECTION

In yet another alternative embodiment, the recombinagenic oligonucleobase can be delivered to the plant cell by whiskers or microinjection of the plant cell. The so called whiskers technique is performed essentially as described in Frame et al., 1994, Plant J. 6:941-948. The recombinagenic oligonucleobase is added to the whiskers and used to transform the plant cells. The recombinagenic oligonucleobase may be co-incubated with plasmids comprising sequences encoding proteins capable of forming recombinase complexes in plant cells such that recombination is catalyzed between the oligonucleotide and the target sequence in the EPSPS gene.

5.4 SELECTION OF GLYPHOSATE RESISTANT PLANTS

Plants or plant cells can be tested for resistance or tolerance to a phosphonomethylglycine herbicide using commonly known methods in the art, e.g., by growing the plant or plant cell in the presence of a herbicide and measuring the rate of growth as compared to the growth rate of control plants in the absence of the herbicide. In the case of glyphosate concentrations of from about 0.01 to about 20 mM are employed in selection medium.

6. EXAMPLE 1: PRODUCTION OF GLYPHOSATE-RESISTANT ARABIDOPSIS EPSPS GENES

The following experiments demonstrate the production of mutant *Arabidopsis thaliana* EPSPS genes which are resistant to the herbicide glyphosate and which allows the plant cells to maintain a growth rate

6.1 MATERIAL AND METHODS

6.1.1 ISOLATION OF ARABIDOPSIS THALIANA EPSPS cDNA

A 1.3 kb DNA fragment was amplified by PCR from an *Arabidopsis* cDNA library using the primers AtEXPEXP1 and AtEXPEXP2CM-2. The two primers were designed to amplify the cDNA from the mature peptide to the termination codon.

WO 03/013226

PCT/US02/21837

The 5' primer AtEXPEXPM1 contains an XbaI site (underlined) and the 3' primer AtEXPEXP2CM-2 contains a BglII site (underlined), sites which will be of use for cloning of the fragment into the expression vector.

AtEXPEXPM1

5'-GCTCTAGAGAAAGCGTCGGAGATTGTACTT-3' (SEQ ID NO:40)

AtEXPEXP2CM-2

5'-GCAGATCTGAGCTCTAGTGCTTGTGATTCTTCAAGTAC-3' (SEQ ID NO:41)

The PCR band was excised from the agarose gel and purified (GeneClean, Biol). Its sequence was then confirmed as the mature peptide sequence of *Arabidopsis thaliana* EPSPS gene.

6.1.2 PREPARATION OF THE EXPRESSION VECTOR

The EPSPS coding region of the *AroE* *Bacillus subtilis* gene was obtained by PCR using the following primers:

BsAroE5'Xba

5'-GCGTCTAGAAAAACGAGATAAGGTGCAG-3' (SEQ ID NO:42) and

BsAroE3'BamHI

5'-GCGGATCCTCAGGATTTTCGAAAGCTTATTTAAATG-3' (SEQ ID NO:43).

The PCR fragment, lacking an initiation codon (ATG), was cloned in-frame to the pACLacIMH6RecA vector by replacing the ORF of *RecA* by digesting with XbaI and BamHI. PACLacIMH6RecA contained the LacI region of Pet21 at positions 1440 to 3176, the MH6 RecA at positions 3809 to 5188, chloramphenicol resistance gene at positions 5445-218 (5446 to 5885 and 1 to 218), and the p15A origin of replication at positions 581 to 1424. The coding region of *RecA* gene was cloned from *E.coli* in-frame with the start codon and 6 histidine linker (MH6) behind the LacZ promoter of pUC19.

WO 03/013226

PCT/US02/21837

6.1.3 CLONING OF THE ARABIDOPSIS EPSPS GENE INTO BACTERIAL EXPRESSION VECTOR

The *Arabidopsis* 1.3 kb PCR fragment was digested with XbaI and BamHI (compatible with BglII) and cloned into the plasmid pACYCLacIMH6EPSPS, in place of the *Bacillus* gene.

The clones obtained (selected on chloramphenicol) were then sequenced and confirmed positive. Confirmed clones are selected and the junctions between the cDNA and the cloning plasmid are confirmed to be identical to the expected sequences.

6.1.4 NOVEL POINT MUTATIONS IN THE EPSPS GENE

Ten different mutants of the *Arabidopsis thaliana* EPSPS gene were designed, (see Figure 2). For the mutagenesis experiments, PCR primers were designed with one, two or three mutations. The PCR reactions are performed using a regular flanking primer (5'ATEPS-198: 5'- GAAAGCGTCGGAGATTGTAC-3') and one of the mutation-carrying primers that correspond to the mutations in Figure 2.

The 353bp PCR fragments obtained are purified (Qiagen PCR Purification kit) and their sequence confirmed. The fragments are then digested with PstI (underlined in the primer sequences) and BamHI and ligated to the pAtEPS-12 vector, which had itself been previously digested with PstI and BamHI.JM109 (Promega) competent cells are used for the transformation and plated onto chloramphenicol-containing LB plates. Clones from each mutagenesis experiment are then isolated and their sequence confirmed.

6.1.5 GLYPHOSATE RESISTANCE ASSAYS

Electrocompetent cells of SA4247, a LacZ - *Salmonella typhi* strain, are prepared according to well known procedures (see Current Protocols in Molecular Biology, (Wiley and Sons, Inc.)). 30 μ l of SA4247 competent cells are electroporated with 20 ng of each plasmid DNA encoding *Arabidopsis* wild-type and mutant EPSPS proteins, *Bacillus* wild-type EPSPS, along with a mock transfection as a control. The settings for electroporation are 25 μ F, 2.5KV and 200 ohms. After electroporation, the cells are transferred into a 15 ml culture tube and supplemented with 970 μ l of SOC medium. The cultures are incubated for 1 $\frac{1}{2}$ hours at 37°C at 225 rpm. 50 μ l of each culture are plated onto LB plates containing 17 μ g/ml chloramphenicol (in duplicates) and incubated overnight at

WO 03/013226

PCT/US02/21837

37°C. On the following day, 5 colonies of each plate are picked and transferred onto M9 plates and incubated overnight at 37°C.

Colonies from the overnight incubation on solid M9 are inoculated into 4 ml of liquid M9 medium and grown overnight at 37°C. On the following day, 25 ml of liquid M9 medium containing chloramphenicol, IPTG and 17 mM or 0 mM Glyphosate (Aldrich, 33775-7) are inoculated with 1-2 ml of each overnight culture (in duplicates), the starting OD (at 600 nm) is measured and all the cultures are normalized to start at the same OD. An OD measurement is taken every hour for seven hours. As a control of the bacterial growth, a culture of untransformed *Salmonella* is also inoculated into plain LB medium.

6.1.7 ISOLATION AND PURIFICATION OF THE EXPRESSED PROTEIN FROM BACTERIAL CLONES

One milliliter of overnight culture of each of the bacterial clones is inoculated into 100 ml of liquid LB medium containing chloramphenicol. The cells are allowed to grow at 37°C until they reach an OD of 0.5-0.7 (approximately 3 ½ hours). IPTG is then added to the cultures to a concentration of 1.0 mM. The cells are grown five additional hours. They are then pelleted at 4000 rpm for 20 minutes at 4°C.

The isolation and the purification of the His-tagged proteins are performed following the Qiagen Ni-NTA Protein Purification System. Cell lysates and eluates are run in duplicates on 12.5% acrylamide gels. One of the gels is silver-stained for immediate visualization, the second gel is transferred onto Millipore Immobilon-P membrane, and blocked overnight in 5% milk in TBS-T. The membrane is then exposed to Anti-His primary antibody solution (Amersham Pharmacia biotech, cat# 37-4710), followed by exposure to Anti-Mouse-IgG secondary antibody solution. (NIF825, from Amersham Pharmacia biotech ECLWestern blotting analysis system, cat# RPN2108). Washes and detection reactions are performed according to the manufacturer instructions. Autoradiograms are developed after 5 minutes exposure.

7. EXAMPLE: Microprojectile Bombardment of a Tobacco (NT-1) Cell Suspension

For microprojectile bombardment of plant cells, the media and protocols found in Gelvin, S.B., et al., (eds) 1991, *Plant Molecular Biology Manual* (Kluwer Acad. Pub.) are followed. Gold particles are coated with a recombinagenic oligonucleobase according the

WO 03/013226

PCT/US02/21837

following protocol. The microprojectiles are first prepared for coating, then immediately coated with the recombinagenic oligonucleobase. To prepare the microprojectiles, suspend 60 mg of gold particles in 1 ml of 100% ethanol. Sonicate the suspension for three, 30 sec bursts to disperse the particles. Centrifuge at 12,000 x g for 30 sec, then discard the supernatant. Add 1 ml of 100% ethanol, vortex for 15 sec, centrifuge at 12,000 x g for 5 min, then discard the supernatant. A 25 μ l suspension of washed gold particles (1.0 μ m diameter, 60 mg/ml) in H₂O is slowly vortexed, then 40 μ l MDON (50 μ g/ml), 75 μ l of 2.5 M CaCl₂, 75 μ l 0.1M spermidine are sequentially added to the suspension. All solutions are ice cold. The completed mixture is vortexed for a further 10 min and the particles are allowed to settle at room temperature for a further 10 min. The pellet is washed in 100% ethanol and resuspended in 50 μ l of absolute ethanol. Biostatic delivery is performed using a Biorad Biostatic gun with the following settings: tank pressure 1100 psi, rupture disks x 2 breaking at 900 psi, particle suspension volume 5 μ l.

Lawns of NT-1 cells of approximately 5 cm in diameter, containing approximately 5 million cells, are grown for three days on standard media at 28°C. Gold particles are coated with a recombinagenic oligonucleobase and shot as above. The cells are cultured a further 2.5 days, suspended and transferred to solid medium supplemented with from about 0.01-20 mM glyphosate for selection of glyphosate-resistant mutant cells.

For more stringent selection of glyphosate-resistant cells, cells are transferred from each bombarded plate to 15 ml tubes containing 5 ml of liquid NT-1 cell suspension medium (CSM: Murashige and Skoog salts [Gibco BRL, Grand Island, NY], 500 mg/l MES, 1 mg/l thiamine, 100 mg/l myoinositol, 180 mg/l KH₂PO₄, 2.21 mg/L 2,4-dichlorophenoxyacetic acid [2,4-D], 30g/L sucrose, pH 5.7) 2 d after bombardment. The tubes are inverted several times to disperse cell clumps. The cells are then transferred to solidified CSM medium (CSM with add 8g/l agar-agar [Sigma, St. Louis, MO]) containing 0.01-20 mM glyphosate. After an appropriate period for selection, actively growing cells (raised, light-colored colonies) are selected and transferred to solidified CSM media containing 0.01-20 mM glyphosate. Three to four weeks later, actively growing cells are selected, then transferred to solidified CSM containing 0.01-20 mM glyphosate. Cells that survive this treatment are then analyzed to determine if they have the mutated EPSPS gene.

WO 03/013226

PCT/US02/21837

8. EXAMPLE: Electroporation of Tobacco Mesophyll Protoplasts

Leaves are harvested from 5- to 6-week-old *in vitro*-grown tobacco plantlets.

For protoplast isolation, the procedure of Gallois et al. (1996, Electroporation of tobacco leaf protoplasts using plasmid DNA or total genomic DNA. Methods in Molecular Biology, Vol. 55: Plant Cell Electroporation and Electrofusion Protocols Edited by: J. A. Nickoloff Humana Press Inc., Totowa, NJ. pp. 89 – 107) is used. The following enzyme solution is used: 1.2 % cellulase R-10 "Onozuka" (Karlan, Santa Rosa, CA), 0.8% macerozyme R-10 (Karlan, Santa Rosa, CA), 90 g/l mannitol, 10 mM MES, filter sterilize, store in 10 ml aliquots at -20°C.

Leaves are cut from the mid-vein out every 1 - 2 mm. They are then placed abaxial side down in contact with 10 ml of enzyme solution in a 100 x 20 mm petri plate. A total of 1 g of leaf tissue is placed in each plate, and the plates are incubated at 25°C in the dark for 16 hr. The digested leaf material is pipetted and sieved through a 100 μ m nylon screen cloth (Small Parts, Inc., Miami Lakes, FL). The filtrate is then transferred to a centrifuge tube and centrifuged at 1,000 rpm for 10 min. All centrifugations for this protocol are performed similarly. The protoplasts collect in a band at the top. The band of protoplasts is then transferred to a clean centrifuge tube to which 10 ml of a washing solution (0.4 M sucrose and 80 mM KCl) is added. The protoplasts are gently resuspended, centrifuged, then washed again. After the last wash, the protoplast density is determined by dispensing a small aliquot onto a hemocytometer.

For electroporation, the protoplasts are resuspended to a density of 1×10^6 protoplasts/ml in electroporation buffer (80 mM KCl, 4 mM CaCl₂, 2mM potassium phosphate, pH 7.2, 8% mannitol). The protoplasts are allowed to incubate at 8°C for 2 hr. After 2 hr, 0.3 ml (3×10^5 protoplasts) are transferred to each 0.4 cm cuvette, then placed on ice. GFP-2 (0.6 - 4 μ g/mL) is added to each cuvette except for an unelectroporated control. The protoplasts are electroporated (250V, capacitance 250 μ F, and time constant 10 - 14 ms). The protoplasts are allowed to recover for 10 min on ice, then transferred to petri plates (100 x 20 mm). After 35 min, 10 ml of POM (80% [v/v] CSM, 0.3M mannitol, 20% [v/v] supernatant from the initial centrifugation of the NT-1 cell suspension prior to protoplast isolation), is added to each plate. The plates are transferred to the dark at 25°C for 24 hr, then

WO 03/013226

PCT/US02/21837

transferred to the light. The protoplast cultures are then maintained according to *Gallois supra*.

9. EXAMPLE: Canola Microspore Isolation, Electroporation, and Embryogenesis

For microspore isolation, canola (*Brassica napus* or *Brassica rapa*) buds of appropriate size (depending on environmental conditions: 12-20°C, 3.5-4.5 mm; 20-23°C, 3.0-3.5 mm; 23-28°C, 2.2-2.8 mm) are picked from approximately 6-10 racemes for a small culture or up to 50 for a large culture. The buds are then placed in a steel sterilization basket. In the hood, buds are sterilized by submersing the sterilization baskets containing the buds into 200 ml of 5.6% bleach for 10 minutes. The sterile buds are then rinsed with 200 ml of cold, sterile water for 5 minutes, twice. The buds are then transferred from the sterilization baskets to a blender cup and 25-30 ml of cold microspore wash (13% sucrose solution, pH 6.0) is added. The buds are homogenized with a blender by alternating high and low speeds, five seconds each, for a total of 20 seconds. (Alternatively, the buds are transferred to the mortar, 30 ml of microspore wash are added, and the tissues are ground up using a pestle for approximately 20 sec.) The contents of the blender cup are poured through nested 63 um and 44 um sterile filters in a beaker-funnel apparatus. The blender cup is then rinsed with 10-15 ml microspore wash. The filtrate is poured into 50 ml plastic centrifuge tubes and the volume is adjusted to 50 ml with microspore wash. The tubes are centrifuged for five minutes at 200 x g. After centrifugation, the dark green supernatant is decanted, leaving a yellow spore pellet at the bottom. The wash procedure is repeated two more times for a total of three centrifugations. The supernatant should become clearer with each wash step. The first two cycles of washing should be done in less than 10 minutes to avoid autotoxicity. After the third spin, the microspores are resuspended in 50ml of NLN liquid culture medium (less NLN can be used, depending on pellet size, to permit an easier volume adjustment after determining initial microspore concentration). To make NLN Medium, combine 0.125 g KNO₃, 1.25 g MgSO₄ 7H₂O, 0.5 g Ca(NO₃)₂ 4H₂O, 0.125 g KH₂PO₄, and 4 ml FeSO₄ EDTA [per 500 ml: 1.39 g FeSO₄ 7H₂O, 1.865 g Na₂ EDTA]. Add 10 ml 100X NN vitamin stock [per L: 0.005 g biotin, 0.05 g folic acid, 0.2 g glycine, 10.0 g myoinositol, 0.5 g nicotinic acid, 0.05 g pyridoxine HCl, 0.05 g thiamine HCl], 10 ml 100X MS micronutrient stock [per L: 2.23 g MnSO₄ 4H₂O, 0.62 g boric acid, 0.86 g ZnSO₄ 7H₂O, 0.025 g Na₂MoO₄ 2H₂O, 0.0025

WO 03/013226

PCT/US02/21837

g CuSO₄ 5H₂O, 0.0025 g CoCl₂·6H₂O], 0.03 g glutathione [reduced form], 0.8 g L-glutamine, 0.1 g L-serine, 130 g sucrose, and adjust the pH to 6.0.

Microspores are electroporated using the protoplast electroporation procedure detailed above for *Brassica napus* or *Brassica rapa*. For *Brassica* or other species, other well-known microspore electroporation protocols can be used, including those provided by manufacturers for use with electroporation equipment, e.g., the Electro Cell Manipulator® (ECM 600, BTX Division of Genetronics) or Electro Square Porator™ (T820, BTX Division of Genetronics).

For example, for *Zea mays*, the following protocol is provided for use with the Electro Square Porator™ (T820, BTX Division of Genetronics). Pollen is collected from greenhouse-grown plants. Supplemental light is provided by high-pressure 400 W sodium lights with an average output of 500 ft-candles to achieve a 16 hr/daylight period. Tassles are shaken the day before electroporation to remove old pollen and to ensure collection of recently mature pollen the next morning. Pollen is germinated for 3-5 minutes before electroporation in 0.20 M sucrose, 1.27 mM Ca(NO₃)₂ 4H₂O, 0.16 mM H₃BO₃, 0.99 mM KNO₃, pH 5.2. The following electroporation settings are used: HV Mode/3 KV, one pulse of 99 µsec pulse length at a voltage of 1.5 kV and field strength of 3.75 kV/cm using a disposable cuvette (p/n 640) with a 4 mm gap. Electroporation is carried out at room temperature using a sample volume of 800 µl.

The following protocol is employed to achieve embryogenesis of the microspores. A hemacytometer is used to determine the microspore concentration at the initial volume by counting all microspores in each of the corner quadrants of the hemacytometer. The new culture is determined using the following equation: (number of cells counted / number of fields counted) (10,000) (initial volume/100,000) = new volume. The required culture density for microspores is between 80,000 and 100,000 spores per ml. The volume of the culture is adjusted accordingly and the culture is mixed well. 15 ml of the culture is pipetted into an appropriate number of petri plates. For even plating, one can make slight adjustments (usually no more than 2-3 ml) to make the culture volume a factor of 15, resulting in even plating. Plates are sealed with a double layer of parafilm and stacked in a 30°C incubator in the dark. After seven days, the plates are observed under an inverted scope

WO 03/013226

PCT/US02/21837

to look for cell divisions and embryo development. If cell divisions and tiny globular embryos are observed, the plates are returned to the incubator for another seven days. Otherwise, the culture is discarded. After 14 days at 30 C, the plates are placed on a shaker at 50 rpm at room temperature in the dark for an additional 14 days. After 28-35 days of culture, embryos should be approximately 5 mm long with visible cotyledons. Embryos are then transferred to solid B5 germination medium and exposed to a temperature of 4°C immediately after transfer to solid medium to increase the yield of mature embryos. To make B5 solid germination medium, combine 400 ml B5 x 10 Stock (per 4 L: 50 g KNO₃, 5 g MgSO₄ 7H₂O, 15 g CaCl₂ 2H₂O, 2.68 g (NH₄)₂SO₄, 3 g NaH₂PO₄ H₂O, 32 ml FeSO₄ EDTA), 200 ml B5 vitamin stock [per L: 10 g myoinositol, 0.1 g nicotinic acid, 0.1 g pyridoxine HCl, 1 g thiamine-HCl], 200 ml 100x B5 micronutrient stock [per L: 1 g MnSO₄ H₂O, 0.3 g H₃BO₃, 0.2 g ZnSO₄ 7H₂O, 0.025 g Na₂MoO₄ 2H₂O, 0.0025 g CuSO₄ 5H₂O, 0.0025 g CoCl₂ 6H₂O], 20 ml KI stock [0.83 g/L KI]; 40 g sucrose; and 2 ml GA₃ stock [0.1 g/L GA]. Bring the volume up to 2 L with double distilled water, pH 5.7, and add 8g agar per L before autoclaving. The embryos are maintained at 4°C for 10 days. The plates are then moved to a light chamber set between 23 and 27°C with a 12 hr light regime. The plates remain in these conditions for 30 days. The plantlets generated after this period can be transferred directly to soil.

The invention claimed and described herein is not to be limited in scope by the specific embodiments herein disclosed since these embodiments are intended as illustrations of several aspects of the invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

A number of references are cited herein, the entire disclosures of which are incorporated herein, in their entirety, by reference.

WO 03/013226

PCT/US02/21837

WE CLAIM:

1. An herbicide resistant plant that expresses a mutant EPSPS gene product wherein the EPSPS gene is mutated at a position to change one or more amino acid positions in the gene product, said amino acid positions selected from the group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Arg₄₈₀, Gly₁₇₇ and Lys₅₀₅ in *Arabidopsis* or at an analogous amino acid position in an EPSPS homolog.

2. The plant according to claim 1 wherein the plant is *Zea mays* and the amino acid positions are selected from the group consisting of Asp₅₁, Gly₁₀₁, Arg₁₃₁, Arg₃₆₂, His₄₀₃, Arg₄₀₄ and Lys₄₂₉.

3. The plant according to claim 1 wherein the plant is *Brassica napus* and the amino acid positions are selected from the group consisting of Asp₁₂₂, Arg₂₀₃, Arg₄₃₄, His₄₇₅, Arg₄₇₆, Gly₁₇₃ and Lys₅₀₁.

4. The plant according to claim 1 wherein the plant is *Petunia hybrida* and the amino acid positions are selected from the group consisting of Asp₁₂₂, Arg₂₀₃, Arg₄₃₄, His₄₇₅, Arg₄₇₆, Gly₁₇₃ and Lys₅₀₁.

5. The plant according to claim 1 wherein the plant is selected from the group consisting of corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar, pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.

6. The plant according to claim 1 in which the mutated gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

- (i) Asp₁₂₆ - Glu
- (ii) Arg₂₀₇ - Glu
- (iii) Arg₄₃₈ - Lys
- (iv) His₄₇₉ - Arg or Leu
- (v) His₄₇₉R₄₈₀ - Arg₄₇₉Lys₄₈₀

WO 03/013226

PCT/US02/21837

(vi) Gly₁₇₇ – Met or Ser

(vii) Lys₅₀₅ – Arg

7. The plant according to claim 2 in which the mutated gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

(i) Asp₅₁ - Glu

(ii) Gly₁₀₁ – Ser or Met

(iii) Arg₁₃₁ - Glu

(iv) Arg₃₆₂ - Lys

(v) His₄₀₃ – Leu or Arg

(vi) His₄₀₃Arg₄₀₄ – Arg₄₀₃Lys₄₀₄

(vii) Lys₄₂₉ - Arg

8. The plant according to claim 3 in which the mutated gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

(i) Asp₁₂₂ - Glu

(ii) Arg₂₀₃ - Glu

(iii) Arg₄₃₄ - Lys

(iv) His₄₇₅ – Leu or Arg

(v) His₄₇₅Arg₄₇₆ – Arg₄₇₅Lys₄₇₆

(vi) Gly₁₇₃ - Met or Ser

(vii) Lys₅₀₁ - Arg.

9. The plant according to claim 4 in which the mutated gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

(i) Asp₁₂₂ - Glu

(ii) Arg₂₀₃ - Glu

(iii) Arg₄₃₄ - Lys

(iv) His₄₇₅ – Leu or Arg

WO 03/013226

PCT/US02/21837

- (v) His₄₇₅Arg₄₇₆ – Arg₄₇₅Lys₄₇₆
- (vi) Gly₁₇₃ - Met or Ser
- (vii) Lys₅₀₁ - Arg.

10. A mutant EPSPS protein comprising the amino acid sequence of the *Arabidopsis* EPSPS gene product depicted in FIG. 1 in which one or more amino acids selected from the group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Gly₁₇₇ and Lys₅₀₅ (or at an analogous amino acid position in an EPSPS homolog) is changed to a different amino acid, which mutant EPSPS protein has increased resistance or tolerance to a phosphonomethylglycine herbicide.

11. The mutant EPSPS protein of Claim 10 further comprising a change at amino acid position Arg₄₈₀ to a different amino acid when amino acid His₄₇₉ is also changed to a different amino acid.

12. The mutant EPSPS protein of Claim 11 wherein His₄₇₉ is changed to Arg₄₇₉ and Arg₄₈₀ is changed to Lys₄₈₀.

13. The mutant EPSPS protein of Claim 10 wherein Asp₁₂₆ is changed to Glu₁₂₆.

14. The mutant EPSPS protein of Claim 10 wherein the Arg₂₀₇ is changed to Glu₂₀₇.

15. The mutant EPSPS protein of Claim 10 wherein the Arg₄₃₈ is changed to Lys₄₃₈.

16. The mutant EPSPS protein of Claim 10 wherein the His₄₇₉ is changed to Leu₄₇₉ or Arg₄₇₉.

17. The mutant EPSPS protein of Claim 10 wherein the Gly₁₇₇ is changed to Ser₁₇₇ or Met₁₇₇.

WO 03/013226

PCT/US02/21837

18. The mutant EPSPS protein of Claim 10 wherein the Lys₅₀₅ is changed to Arg₅₀₅.

19. A method for producing an herbicide resistant or tolerant plant which comprises:

- a. introducing into a plant cell a recombinogenic oligonucleobase to produce a mutant EPSPS gene wherein the EPSPS gene is mutated at a position to change one or more amino acid positions in the gene product, said amino acid positions selected from the group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Arg₄₈₀, Gly₁₇₇ and Lys₅₀₅ in *Arabidopsis* or at an analogous amino acid position in an EPSPS homolog.; and
- b. identifying a cell having a mutated EPSPS gene.

20. The method of Claim 19 wherein the mutated EPSPS gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

- (i) Asp₁₂₆ - Glu
- (ii) Arg₂₀₇ - Glu
- (iii) Arg₄₃₈ - Lys
- (iv) His₄₇₉ - Arg or Leu
- (v) His₄₇₉R₄₈₀ - Arg₄₇₉Lys₄₈₀
- (vi) Gly₁₇₇ - Met or Ser
- (vii) Lys₅₀₅ - Arg.

21. The method of Claim 19 wherein plant is a *Zea mays* plant and the amino acid positions in the *Zea mays* homolog are selected from the group consisting of Asp₅₁, Gly₁₀₁, Arg₁₃₁, Arg₃₆₂, His₄₀₃, Arg₄₀₄ and Lys₄₂₉.

22. The method of Claim 21 wherein the mutated EPSPS gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

- (i) Asp₅₁ - Glu

WO 03/013226

PCT/US02/21837

- (ii) Gly₁₀₁ – Ser or Met
- (iii) Arg₁₃₁ - Glu
- (iv) Arg₃₆₂- Lys
- (v) His₄₀₃ – Leu or Arg
- (vi) His₄₀₃Arg₄₀₄ – Arg₄₀₃Lys₄₀₄
- (vii) Lys₄₂₉ – Arg.

23. The method of Claim 19 wherein the plant is a *Brassica napus* plant and the amino acid positions in the *Brassica napus* homolog are selected from the group consisting of Asp₁₂₂, Arg₂₀₃, Arg₄₃₄, His₄₇₅, Arg₄₇₆, Gly₁₇₃ and Lys₅₀₁.

24. The method of Claim 23 wherein the mutated EPSPS gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

- (i) Asp₁₂₂ - Glu
- (ii) Arg₂₀₃ - Glu
- (iii) Arg₄₃₄ - Lys
- (iv) His₄₇₅ – Leu or Arg
- (v) His₄₇₅Arg₄₇₆ – Arg₄₇₅Lys₄₇₆
- (vi) Gly₁₇₃ - Met or Ser
- (vii) Lys₅₀₁ – Arg.

25. The method of Claim 19 wherein the plant is a *Petunia hybrida* plant and the amino acid positions in the *Petunia hybrida* are selected from the group consisting of Asp₁₂₂, Arg₂₀₃, Arg₄₃₄, His₄₇₅, Arg₄₇₆, Gly₁₇₃ and Lys₅₀₁.

26. The method of Claim 25 wherein the mutated EPSPS gene results in one or more of the following amino acid substitutions in the EPSPS gene product in comparison with the wild-type sequence:

- (i) Asp₁₂₂ - Glu
- (ii) Arg₂₀₃ - Glu
- (iii) Arg₄₃₄ - Lys
- (iv) His₄₇₅ – Leu or Arg

WO 03/013226

PCT/US02/21837

- (v) His₄₇₅Arg₄₇₆ – Arg₄₇₅Lys₄₇₆
- (vi) Gly₁₇₃ – Met or Ser
- (vii) Lys₅₀₁ – Arg.

27. The method of Claim 19 wherein the recombinagenic oligonucleobase is a mixed duplex nucleotide or a SSMOV.

28. The method of Claim 27 wherein the mixed duplex nucleotide contains a first homologous region which has a sequence identical to the sequence of at least 6 base pairs of the first fragment of the target EPSPS gene and a second homologous region which has a sequence identical to the sequence of at least 6 based pairs of a second fragment of the target EPSPS gene, and an intervening region which contains at least one nucleobase heterologous to the target EPSPS gene, which intervening region connects the first and second homologous region.

29. The method of Claim 19 wherein the recombinagenic oligonucleobase is introduced by electroporation.

30. The method of Claim 19 in which the plant is selected from the group consisting of the plant may be selected from a species of plant from the group consisting of canola, sunflower, tobacco, sugar beet, cotton, maize, wheat, barley, rice, sorghum, tomato, mango, peach, apple, pear, strawberry, banana, melon, potato, sweet potato, yam, carrot, lettuce, onion, soya spp, sugar cane, pea, peanut, field beans, poplar, grape, citrus, alfalfa, rye, oats, turf grasses, forage grasses, flax, oilseed rape, cucumber, morning glory, balsam, pepper, eggplant, marigold, lotus, cabbage, daisy, carnation, tulip, iris, lily, nut producing plants, pine, eucalyptus, lentils, and other *Brassica* sp.

31. A method of making a glyphosate resistant plant which comprises:

a. providing a recombinagenic oligonucleobase to produce a mutant EPSPS gene wherein the EPSPS gene is mutated at a position to change one or more amino acid positions in the gene product, said amino acid positions selected from the

WO 03/013226

PCT/US02/21837

group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Arg₄₈₀, Gly₁₇₇ and Lys₅₀₅ in *Arabidopsis* or at an analogous amino acid position in an EPSPS homolog;

- b. introducing said recombinagenic oligonucleotide into a plant cell;
- c. culturing said cell to obtain descendant plant cells, said descendant plant cells containing the mutant EPSPS gene; and
- d. establishing that the mutant EPSPS gene is expressed in said descendant plant cells.

32. A method of making seeds that will grow into plants that are resistant to glyphosate herbicide which comprises:

- a. providing a recombinagenic oligonucleobase to produce a mutant EPSPS gene wherein the EPSPS gene is mutated at a position to change one or more amino acid positions in the gene product, said amino acid positions selected from the group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Arg₄₈₀, Gly₁₇₇ and Lys₅₀₅ in *Arabidopsis* or at an analogous amino acid position in an EPSPS homolog;
- b. introducing said recombinagenic oligonucleotide into a plant cell;
- c. culturing said cell to obtain descendant plant cells, said descendant plant cells containing the mutant EPSPS gene; and
- d. establishing that the mutant EPSPS gene is expressed in said descendant plant cells
- e. regenerating a whole fertile plant that expresses the mutant EPSPS gene; and
- f. collecting the seed from the whole fertile plant.

WO 03/013226

PCT/US02/21837

33. The method of Claim 32 wherein the seed is germinated to produce more seed containing the mutant EPSPS gene and glyphosate is applied to the germinated plants to kill any plants that do not contain the mutated EPSPS gene.

34. A method of selectively cultivating EPSPS mutant plants which comprises:

a. cultivating EPSPS mutant plants wherein the EPSPS gene is mutated at a position to change one or more amino acid positions in the gene product, said amino acid positions selected from the group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Arg₄₈₀, Gly₁₇₇ and Lys₅₀₅ in *Arabidopsis* or at an analogous amino acid position in an EPSPS homolog;

b. applying a sufficient amount of glyphosate herbicide to the cultivated mutant plants of (a) such that the glyphosate is toxic to at least one non-mutant plant.

35. A method of propagating an EPSPS mutant plant wherein the EPSPS gene is mutated at a position to change one or more amino acid positions in the gene product, said amino acid positions selected from the group consisting of Asp₁₂₆, Arg₂₀₇, Arg₄₃₈, His₄₇₉, Arg₄₈₀, Gly₁₇₇ and Lys₅₀₅ in *Arabidopsis* or at an analogous amino acid position in an EPSPS homolog which comprises (1) vegetatively propagating a plant containing said EPSPS mutation or (2) culturing a plant cell or plant tissue containing said EPSPS mutation to form callus tissue and regenerating a plant therefrom wherein the regenerated plant contains said EPSPS mutation.

WO 03/013226

PCT/US02/21837

FIGURE 1
(page 1)

+1	M A S S	L T S	K S I	L G C T	K P A
1	ATGGCGTCTT	CTCTCACTTC	CAAATCCATT	CTCGGATGCA	CCAAACCCGC
	TACCGCAGAA	GAGAGTGAAG	GTITAGGTAA	GAGCCTACGT	GGTTTGGCG
+1	A 6 S S F I P S E L R R L S S S P A V				
51	TTCCTCTCT	TTTCTCCGT	CGGAGCTCCG	TCGTCTCTCT	TCTCCCGCCG
	AAGAAGAAGA	AAAGAAGGCA	GCCTCGAGGC	AGCAGAGAGA	AGAGGGCGGC
+1	V Q I S L H S Q T R K N F R Q S W				
101	TTCAGATATC	TCTCCATTCA	CAAACCCAGGA	AGAACTTCCG	GCAGTCGTGG
	AAGTCTATAG	AGAGGTAAGT	GTITGGTCT	TCTTGAAGGC	CGTCACGACC
+1	G L K K S D L M L N G S E I R P V				
151	GGATGAAAGA	AGRGTAATCT	GATGCTAAT	GGTTCTGAGA	TTCGTCCTGT
	CCTAACTCT	TCTCACTAGA	CTACGATTTA	CCAGACTCT	AAGCAGGACA
+1	V K V R A S V S T A E K A S E I V L				
201	GAAGGTTAGG	GCTCTGTTT	CCACGGCGGA	GAAGGCTCG	GAGATTGTC
	CTTCCTATCC	CGAAGACAAA	GGTGGCCCT	CTTCGAGAC	CTCTACACG
+1	L Q P I R E I S G L I K L P G S K				
251	TTCAACCCAT	TAGAGAATTC	TCGGGTCTCA	TTAAGCTTCC	TGGCTCCAAG
	AAGTGGGTA	ATCTCTTGTAG	AGCCAGAGT	AATTGAAAGG	ACCGAGGTTC
+1	S L S N R I L L A A L S E G T T				
301	TCTCTCTCTA	ATCGAATTCT	GCTCTCGCT	GCTCTATCTG	AGGGAACTAC
	AGAGAGAGAT	TAGCTTAAGA	CGAAGAGCGA	CGAGATAGAC	TCCCTGTATG
+1	T V V D N L L N S D D I N Y M L D A				
351	TGTAGTGGAC	AACTTGTGTA	ACAGTGTGTA	CATCAATTAC	ATGCTGTG
	ACATCACCTG	TTGAAACAAT	TGTCACTACT	GTAGTTATG	TACGAACTAC
+1	A L K I I L G L N V E T H S E N H R				
401	CGTGAAGAT	ATGGGACTT	ATATGTGGAAA	CTCACAGTGA	AAACAACTCGT
	GCACATTCTA	TAACCTGTAA	TTACACCTTT	GAGTGTCACT	TTTGTGTGCA
+1	A V V E G C G G G V F P A S I D S K				
451	GCTGTAGTTG	AAAGGATGTGG	CGGGGTATTT	CCAGCTTCCA	TTGATTCCA
	CGACATCAC	TTCCTACACC	CCCCCATAAA	GGTCGAAGGT	AACTAAGGTT
+1	K S D I E L Y L G N A G T A M R P L				
501	GAGTGTATC	GAACCTTAC	TGGGCAATGCG	AGGAACACCA	ATGGTCCAC
	CTCACTATAG	CTTGAATGG	AGCCGTTACG	TCTTGTGCGT	TACCGAGGTG
+1	L T A A V T A A G G N A S Y V L D				
551	TTACCGCCG	AGTTACTGCT	GCAGGTGGCA	ACGCAAGTTA	TGTCCCTGTAT
	AATGGGGGG	TCAATGACGA	CGTCCACCGT	TGCGTTCAAT	ACAGGAACTA
+1	G V P R M R E R P I G D L V V G L				
601	GGGGTGCCTC	GGATGAGAGA	GAGACCTATA	GGGGATTGG	TTGTTGGCT
	CCCCACGGAG	CCTACTCTCT	CTCTGGATAT	CCCCTAAACC	AAACACCGAA
+1	L K Q L G A D V E C T L G T N C P P				
651	TAAGCAGCTT	GGTGTGATG	TTGAATGTAC	TCTTGGCACT	AACTGCCCTC
	ATTGTCGAA	CCACGACTAC	AACTTACATG	AGAACCGTGA	TTGACGGGAG
+1	P V R V N A H G G L P G G K V K L				
701	CTGTGCGT	CAACGCTAAT	GGTGGCCCTTC	CTGGTGGATA	GGTGAAGCTT
	GACAAGCACA	GTGCGGATTA	CCACCGGAAG	GACCACCTTT	CCACTTCGAA
+1	S G S I S S Q Y L T A L L M A A P				
751	TCTGGATCTA	TTAGTAGTCA	GTACTTGACC	GCTCTGCTCA	TGGCACCTCC
	AGACCTAGAT	AATCATCGAT	CATGAACCTGG	CGAGACGAGT	ACCGTCGAGG

WO 03/013226

PCT/US02/21837

FIGURE 1
(page 2)

+1	P L A L G D V E I E I V D K L I S V
801	CTTAGCTCTT GGAGACGTCG AAATTGAAAT TGTCGATAAA TTGATTTCTG GAATCGAGAA CCTCTCGACG TTTAACCTTA ACAGCTAATT AACTAAAGAC
+1	V P Y V E M T L K L M E R F G V S
851	TTCCGTATGT TGAAAATGACA TTGAAAGTGA TGGAAACGTTT TGGGGTAAGT AAGGCATACA ACTTTACTGT AACTTCAACT ACCTTGCAAA ACCCCATTCA
+1	A E H S E S W D R F F V K G G Q K
901	GCTGAGCATA GTGAAAGCTG GGATCGTTTC TTGTTGTAAGG GTGGGGAAAA CGACTCGTAT CACTTTGAC CCTAGCAAG AAACATTCC CACCCGTTTT
+1	K Y K S P G N A Y V E G O A S S A S
951	ATACAAGTCG CCGGGTAATG CTTACGTAGA AGGTGATGCT TCTAGTGCTA TATGTTCAAGC GGCCCATTAC GATGCACT TCCACTACGA AGATCACGAT
+1	S Y F L A G A A I T G E T V T V E
1001	GTATTTCTT GGCTGGTGCT GCCATTACCG GTGAAACTGT CACTGTTGA CAATAAAGGA CCGACCACGA CGGTAATGGC CACTTGACA GTGACAACCT
+1	G C O T T S L O G D V K F A E V L
1051	GGTTGTGGAA CGACCACTTT GCAGGGAGAT GTGAAATTTG CCGAGGTCT CGAACACCTT GCTGGTCAAA CGTCCCTCTA CACTTAAC GGCTCCAAGA
+1	L E K M G C K V S W T E N S V T V T
1101	TGAGAAAATG GGATGTAAG TGTCCTGGAC AGAGAACAGT GTGACTGTGA ACTCTTTAC CCTACATTTC ACAGGACTG TCTCTGTCA CACTGACACT
+1	T G P S R D A F G M R H L R A I D
1151	CAGGGCCGTC TAGAGATGCT TTGGAAATGA GACACCTTGGC GGCTATTGAT GTCCCGGCAG ATCTCTACCA AAACCTTAAC CTGTGAACGC CCGATAACTA
+1	V N M N K M P D V A M T L A V V A
1201	GTCAACATGA ACAAAATGCC TGATGTAGCA ATGACTCTTG CCGTCGTTGC CAGTTGTACT TGTTTTACGG ACTACATCGT TACTGAGAAC GGCAGCAACG
+1	A L F A D G P T T I R D V A S W R V
1251	TCTCTTGCC GATGGTCCAA CCACCATAG AGATGTGGCT AGCTGGAGAG AGAGAAAACGG CTACCAAGTTT GGTGGTAATC TCTACACCGA TCGACCTCTC
+1	V K E T E R M I A I C T E L R K L
1301	TAAGGGAGAC GGAAAGGATG ATTGCCATT GCACAGAGCT TAGAAAACGT ATTTCCCTGC CTTTCTCTAC TAACGGTAAA CGTGTCTCGA ATCTTTGAC
+1	G A T V E E G S D Y O V I T P P K
1351	GGAGCTACAG TGGAAAGARGG TTCAAGTTAT TGTTGATTA CTCCGCCGAA CTTCGATGTC ACCTTCTTCC AAGTCTAATA ACACACTAAT GAGGGGGCTT
+1	K K V K P A E I D T Y D D H R M A M
1401	AAAGGTGAAA CGGGCAGAGA TTGATACATA TGATGATCAT AGAATGGCAA TTTCCACTTT GGCCGTCTCT AACTATGTAT ACTACTAGTA TCTACCGTT
+1	M A F S L A A C A D V P I T I N D
1451	TGGCATTCTC TCTTGCAAGCT TGTGCTGATG TTCCAAATCAC CATCAATGAC ACCGTAAGAG AGAACGTCGA ACACGACTAC ARGGTTAGTG GTAGTTACTG
+1	P G C T R K T F P D Y F Q V L E R
1501	CCGGTTGCA CCAGGAAAC CTTCCCCGAC TACTTCCAAG TCCTTGAAAG GGGCCAACGT GGTCCCTTTTG GAAGGGGCTG ATGAGGGTTC AGGAACCTTC
+1	R I T K H *
1551	AATCACAAAG CATTAG TTAGTGTTC GAAATC

WO 03/013226

PCT/US02/21837

FIGURE 2
(page 1)

1. LIST OF NEW MUTANTS

10 to 14 changes 1 to 5 in a G₉₆ → A₉₆ background

	ECOLI	ARABIDOPSIS	MUTATION	
1.	D ₄₉ → E ₄₉	D126E	GAC → GAA or GAG	}
2.	R ₁₂₄ → K ₁₂₄	R207K	AGA → AAA	
3.	R ₃₄₄ → K ₃₄₄	R438K	AGG → AAG	
4.	H ₃₈₅ → L ₃₈₅	H479L	CAT → CTT	
5.	H ₃₈₅ → R ₃₈₅	H479R	CAT → CGT	
6.	H ₃₈₅ R ₃₈₆ → R ₃₈₅ K ₃₈₆	HR479480RK	(R → K) AGA → AAA	
7.	G ₉₆ → S ₉₆	G177S	GGA → TCA	
8.	G ₉₆ → M ₉₆	G177M	GGA → ATG	
9.	K ₄₁₁ → R ₄₁₁	K505R	AAA → AGA	

1 → 9 on wild type background

1 → 5 on G → A177 background too

a. D126E

	ECOLI	ARABIDOPSIS	MUTATION
1.	D ₄₉ → E ₄₉	D126E	GAC → GAG

ACAACTTGTGAATAGCGATGACATCAATTACATGCTTGATGCG
ACAACTTGTGAATAGCGATGAGATCAATTACATGCTTGATGCG

b. R207K

	ECOLI	ARABIDOPSIS	MUTATION
2.	R ₁₂₄ → K ₁₂₄	R207K	AGA → AAA

GGGGTGCCTCGTATGAGAGAAAAGACCTATAAGGGGATTTGGTTGG
GGGGTGCCTCGTATGAGAGAAAAGACCTATAAGGGGATTTGGTTGG

c. R438K

3.	R ₃₄₄ → K ₃₄₄	R438K	AGG → AAG
----	-------------------------------------	-------	-----------

TGGAGAGTAAAGGAGACAGAAAGGATGATTGCCATTGCACAGA
TGGAGAGTAAAGGAGACAGAAAAGATGATTGCCATTGCACAGA

d. H479L

H ₃₈₅ → L ₃₈₅	H479L	CAT → CTT
-------------------------------------	-------	-----------

CAGAGATTGATACATATGATGATCATAGAATGGCAATGGCATTCTCT
CAGAGATTGATACATATGATGATCAGAATGGCAATGGCATTCTCT

WO 03/013226

PCT/US02/21837

FIGURE 2
(page 2)

e. H479R

H ₃₈₃ → R ₃₈₅	H479R	CAT → CGT
-------------------------------------	-------	-----------

GAGATTGATACATATGATGATCATAGAATGGCAATGGCATTCTCTC
GAGATTGATACATATGATGATC_gTAGAATGGCAATGGCATTCTCTC

f. HR479480RK

H ₃₈₃ R ₃₈₆ → R ₃₈₅ K ₃₈₆	HR479480RK	(R → K) AGA → AAA
---	------------	-------------------

GAGATTGATACATATGATGATCATAGAATGGCAATGGCATTCTCTC
GAGATTGATACATATGATGATC_gTAaAATGGCAATGGCATTCTCTC

g. K505R

K ₄₁₁ → R ₄₁₁	K505R	AAA → AGA
-------------------------------------	-------	-----------

CAACGACTCTGGTTGCACCAAGGAAAACCTTCCCCGACTACTTCCAA
CAACGACTCTGGTTGCACCAAGGAgAACCTTCCCCGACTACTTCCAA

h. G177S

G ₉₆ → S ₉₆	G177S	GGA → TCA
-----------------------------------	-------	-----------

TATCGAACTTTACCTCGGTAAATGCAGGAACAGCAATGCGTCCACTTACCGC
TATCGAACTTTACCTCGGTAAATGCAtAACAGCAATGCGTCCACTTACCGC

i. G177M

G ₉₆ → M ₉₆	G177M	GGA → ATG
-----------------------------------	-------	-----------

TATCGAACTTTACCTCGGTAAATGCAGGAACAGCAATGCGTCCACTTACCGC
TATCGAACTTTACCTCGGTAAATGCAatgACAGCAATGCGTCCACTTACCGC

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

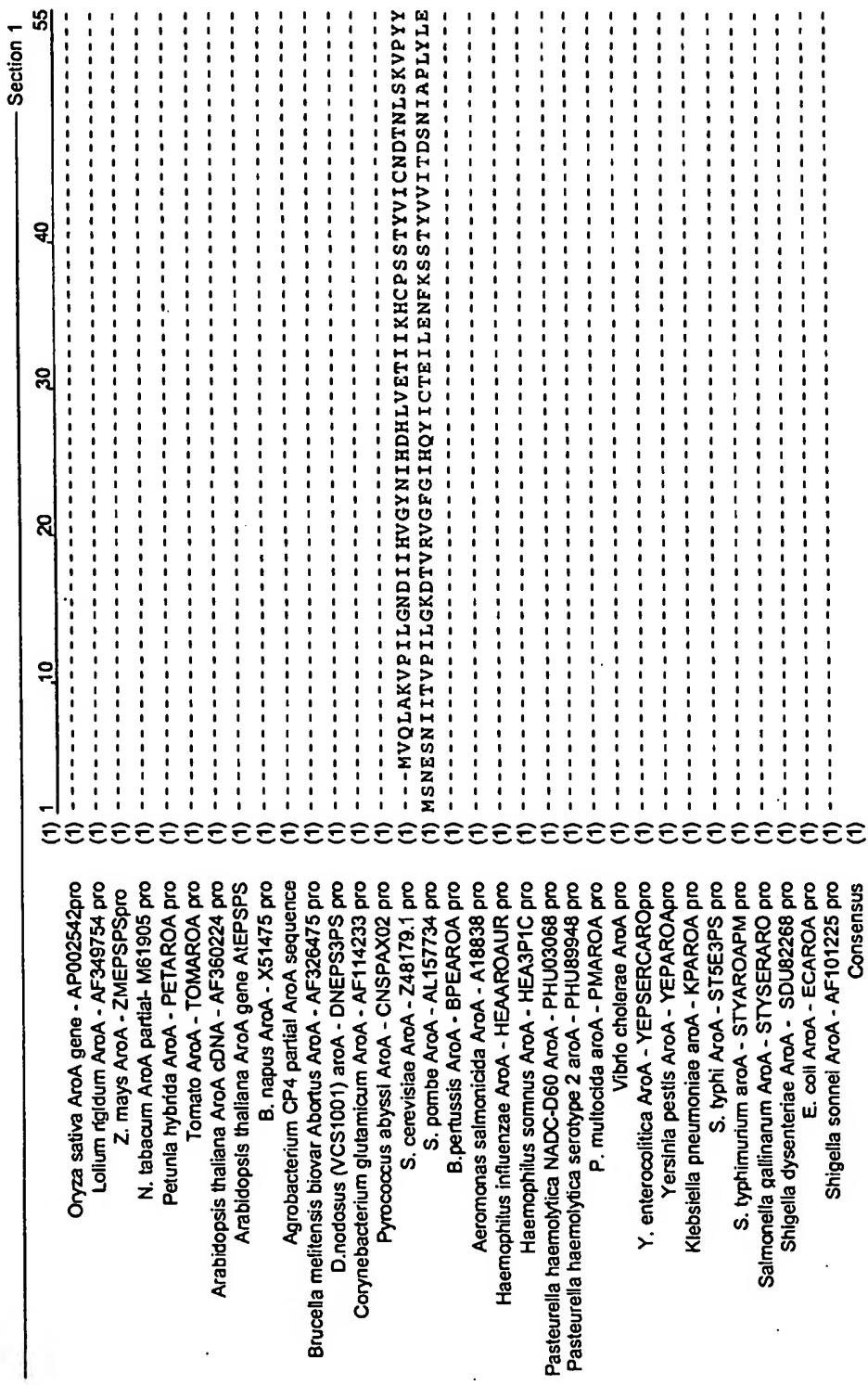


FIGURE 3
Page 1

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

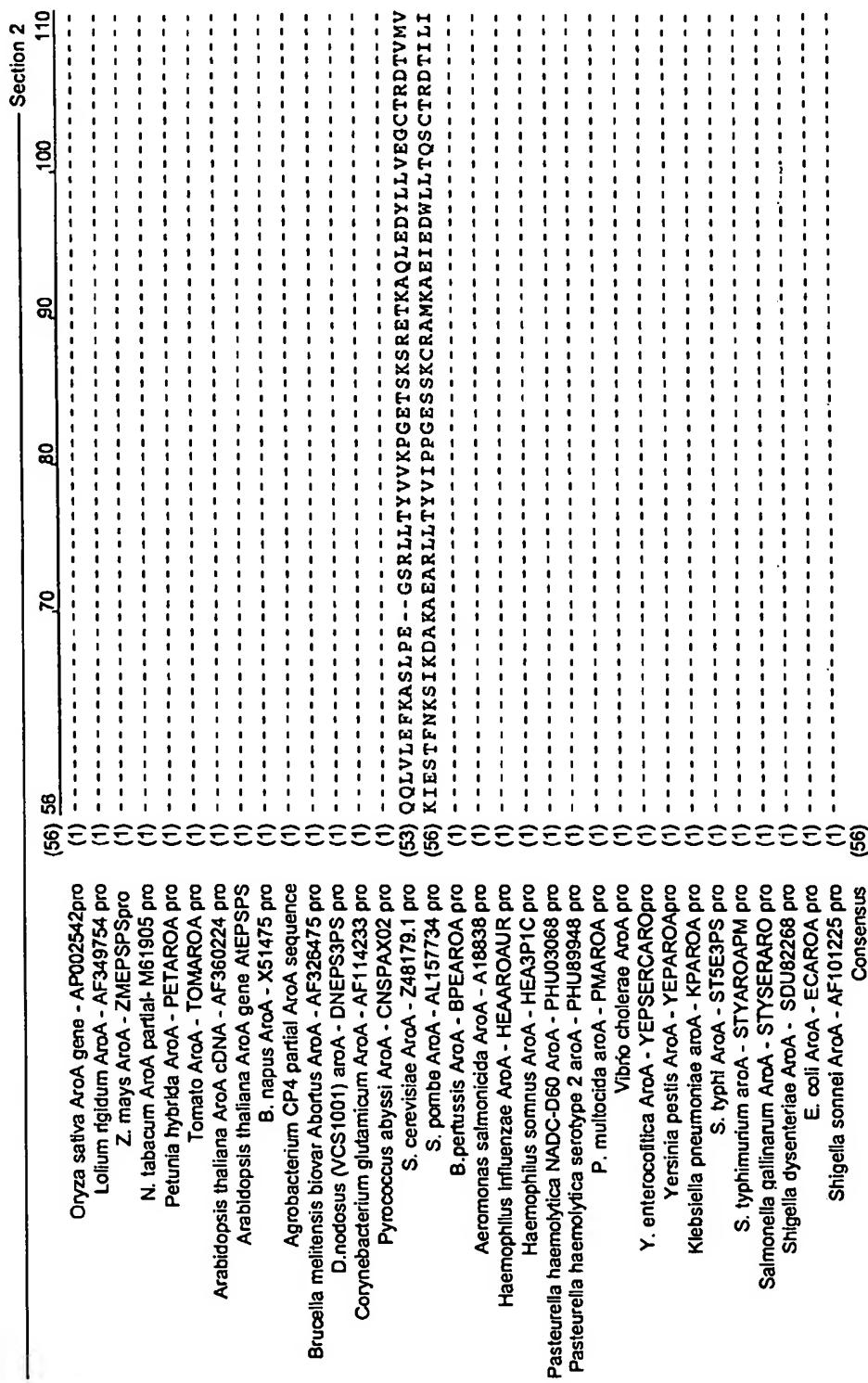


FIGURE 3
Page 2

EPSP Synthase CDS protein alignment

	Section 3				
	111	120	130	140	150
<i>Oryza sativa</i> AroA gene - AP002542 pro	(1)	-	-	-	-
<i>Lolium rigidum</i> AroA - AF349754 pro	(1)	-	-	-	-
<i>Z. mays</i> AroA - ZMEPSPSpro	(1)	-	-	-	-
<i>N. tabacum</i> AroA partial- M61905 pro	(1)	-	-	-	-
<i>Petunia hybrida</i> AroA - PETAROA pro	(1)	-	-	-	-
<i>Tomato</i> AroA - TOMAROA pro	(1)	-	-	-	-
<i>Arabidopsis thaliana</i> AroA cDNA - AF360224 pro	(1)	-	-	-	-
<i>Ara bidopsis</i> , <i>thaliana</i> AroA gene AtEPSPS	(1)	-	-	-	-
<i>B. napus</i> AroA - X51475 pro	(1)	-	-	-	-
<i>Agrobacterium</i> CP4 partial AroA sequence	(1)	-	-	-	-
<i>Brucella melitensis</i> biovar <i>Abortus</i> AroA - AF326475 pro	(1)	-	-	-	-
<i>D. nodosus</i> (VCS1001) aroA - DNEPS3PS pro	(1)	-	-	-	-
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(1)	-	-	-	-
<i>Pynococcus abyssi</i> AroA - CNSPAX02 pro	(1)	-	-	-	-
<i>S. cerevisiae</i> AroA - Z48179.1 pro	(106)	AIGGGVIGDMIGFVASTFMRGVVRVVQVPTSSLAMVDSIGGGKTAIDTPLGKNFIG	-	-	-
<i>S. pombe</i> AroA - ALI57734 pro	(111)	AMGGGVIGDLDLVGYVAASFMRGIRFIQMPPTLLAMVDSIGGGKTDIDTPLGKNLVG	-	-	-
<i>B. pertussis</i> AroA - BPEAROA pro	(1)	-	-	-	-
<i>Aeromonas salmonicida</i> AroA - A18838 pro	(1)	-	-	-	-
<i>Haemophilus influenzae</i> AroA - HEAAROAUR pro	(1)	-	-	-	-
<i>Haemophilus somnis</i> AroA - HEA3P1C pro	(1)	-	-	-	-
<i>Pasteurella haemolytica</i> NADC-D60 AroA - PHU03068 pro	(1)	-	-	-	-
<i>Pasteurella haemolytica</i> serotype 2 aroA - PHU89948 pro	(1)	-	-	-	-
<i>P. multocida</i> aroA - PMAROA pro	(1)	-	-	-	-
<i>Vibrio cholerae</i> AroA pro	(1)	-	-	-	-
<i>Y. enterocolitica</i> AroA - YEPSERCARO pro	(1)	-	-	-	-
<i>Yersinia pestis</i> AroA - YEPAROA pro	(1)	-	-	-	-
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(1)	-	-	-	-
<i>S. typhi</i> AroA - ST5E3PS pro	(1)	-	-	-	-
<i>S. typhimurium</i> aroA - STYAROA PM pro	(1)	-	-	-	-
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(1)	-	-	-	-
<i>Shigella dysenteriae</i> AroA - SDU82268 pro	(1)	-	-	-	-
<i>E. coli</i> AroA - ECAROA pro	(1)	-	-	-	-
<i>Shigella sonnei</i> AroA - AF101225 pro	(1)	-	-	-	-
Consensus	(11)	-	-	-	-

FIGURE 3
Page 3

WO 03/013226

PCT/US02/21837

EPSP synthase CDS protein alignment

	Section 4				
	166	180	200	210	220
Oriza saliva AroA gene - AP002542 pro	(1)				
Lolium rigidum AroA - AF349754 pro	(1)				
Z. mays AroA - ZMEPSPS pro	(1)				
N. tabacum AroA partial - M61905 pro	(1)				
Petunia hybrida AroA - PETAROA pro	(1)				
Tomato AroA - TOMAROA pro	(1)				
Arabidopsis thaliana AroA cDNA - AF360224 pro	(1)				
Arabidopsis thaliana AroA gene AtEPSPS	(1)				
B. napus AroA - X51475 pro	(1)				
Agrobacterium CP4 partial AroA sequence	(1)				
Bacillus meijensis biovar Abortus AroA - AF328475 pro	(1)				
D.nodosus (VCS1001) aroA - DNEPSS3PS pro	(1)				
Corynebacterium glutamicum AroA - AF114233 pro	(1)				
Pyrococcus abyssi AroA - CNSPAX02 pro	(1)				
S. cerevisiae AroA - Z4B179.1 pro	(161)	AFWQPKFVLVDIKWLETLAKREFINGMAEVVIKTAIWNNADEFTRILESNASLFLNV			
S. pombe AroA - AL157734 pro	(166)	AFWQPLRYYVDMVFLHTLPPQVINGLSEIIKTAAMWNENDFQLENNSAVLDAA			
B.pertussis AroA - BPEAROA pro	(1)				
Aeromonas salmonicida AroA - A18838 pro	(1)				
Haemophilus influenzae AroA - HEAAROAUR pro	(1)				
Haemophilus somnis AroA - HEA3P1C pro	(1)				
Pasteurella haemolytica NADC-D60 AroA - PHU03058 pro	(1)				
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(1)				
P. multocida aroA - PMAROA pro	(1)				
Vibrio cholerae AroA pro	(1)				
Y. enterocolitica AroA - YEPSERCARO pro	(1)				
Yersinia pestis AroA - YEPAROA pro	(1)				
Klebsiella pneumoniae aroA - KPAROA pro	(1)				
S. typhi AroA - ST5E3PS pro	(1)				
S. typhimurium aroA - STYAROAPM pro	(1)				
Salmonella gallinarum AroA - STYSERARO pro	(1)				
Shigella dysenteriae AroA - SDUB2268 pro	(1)				
E. coli AroA - ECAROA pro	(1)				
Shigella sonnei AroA - AF101225 pro	(1)				
Consensus	(166)				

EPSP Synthase CDS protein alignment

	Section 5			
	221	230	240	250
(221) <i>Oryza sativa</i> AroA gene - AP002542 pro	(1)			
<i>Lolium rigidum</i> AroA - AF349754 pro	(1)			
<i>Z. mays</i> AroA - ZMEPSPSP pro	(1)			
<i>N. tabacum</i> AroA partial- M61905 pro	(1)			
<i>Petunia hybrida</i> AroA - PETAROA pro	(1)			
<i>Tomato</i> AroA - TOMAROA pro	(1)			
<i>Arabidopsis thaliana</i> AroA cDNA - AF360224 pro	(1)			
<i>Arabidopsis thaliana</i> AroA gene ATEPSPS	(1)			
<i>B. napus</i> AroA - X51475 pro	(1)			
<i>Agrobacterium</i> CP4 partial AroA sequence	(1)			
<i>Brucella melitensis</i> biovar <i>Abortus</i> AroA - AF326475 pro	(1)			
<i>D. nodosus</i> (VCS1001) AroA - DNEPSSPS pro	(1)			
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(1)			
<i>Pyrococcus abyssi</i> AroA - CNSPA02 pro	(1)			
<i>S. cerevisiae</i> AroA - Z48179.1 pro	(216)	VNGAXNVXVTNQLTNEIDEISNTDIEAMLDTHTYKIVLESIRVKAEVVSSDERESS		
<i>S. pombe</i> AroA - AL157734 pro	(221)	LN-----KP-----SVPGEYKFDISKPLQKIISSIRTKCEVVTLDEHEGG		
<i>B. pertussis</i> AroA - BPEAROA pro	(1)			
<i>Aeromonas salmonicida</i> AroA - A18838 pro	(1)			
<i>Haemophilus influenzae</i> AroA - HEAAROAUR pro	(1)			
<i>Haemophilus somnis</i> AroA - HFA3P1C pro	(1)			
<i>Pasteurella haemolytica</i> NADC-D60 AroA - PHU03098 pro	(1)			
<i>Pasteurella haemolytica</i> serotype 2 aroA - PHU89948 pro	(1)			
<i>P. multocida</i> aroA - PMAROA pro	(1)			
<i>Vibrio cholerae</i> AroA pro	(1)			
<i>Y. enterocolitica</i> AroA - YEPSERCARO pro	(1)			
<i>Yersinia pestis</i> AroA - YEPAROA pro	(1)			
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(1)			
<i>S. typhi</i> AroA - ST5E3PS pro	(1)			
<i>S. typhimurium</i> aroA - STYAROAPM pro	(1)			
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(1)			
<i>Shigella dysenteriae</i> AroA - SDU82268 pro	(1)			
<i>E. coli</i> AroA - ECAROA pro	(1)			
<i>Shigella sonnei</i> AroA - AF101225 pro	(1)			
<i>Consensus</i>	(221)			

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

	(276)	278	290	292	310	320	330	Section 6
Oryza sativa AroA gene - AP002542 pro	(1)	-----	-----	-----	-----	-----	-----	
Lolium rigidum AroA - AF349754 pro	(1)	-----	-----	-----	-----	-----	-----	
Z. mays AroA - ZMEPSPSpro	(1)	-----	-----	-----	-----	-----	-----	
N. tabacum AroA partial- M61905 pro	(1)	-----	-----	-----	-----	-----	-----	
Petunia hybrida AroA - PETAROA pro	(1)	-----	-----	-----	-----	-----	-----	
Tomato AroA - TOMAROA pro	(1)	-----	-----	-----	-----	-----	-----	
Arabidopsis thaliana AroA cDNA - AF360224 pro	(1)	-----	-----	-----	-----	-----	-----	
Arabidopsis thaliana AroA gene AtEPSPS	(1)	-----	-----	-----	-----	-----	-----	
B. napus AroA - X51475 pro	(1)	-----	-----	-----	-----	-----	-----	
Agrobacterium CP4 partial AroA sequence	(1)	-----	-----	-----	-----	-----	-----	
Brucella melitensis biovar Abortus AroA - AF326475 pro	(1)	-----	-----	-----	-----	-----	-----	
D.nodosus (VCS1001) aroA - DNEPS3PS pro	(1)	-----	-----	-----	-----	-----	-----	
Corynebacterium glutamicum AroA - AF114233 pro	(1)	-----	-----	-----	-----	-----	-----	M
Pyrococcus abyssi AroA - CNSPAX02 pro	(1)	-----	-----	-----	-----	-----	-----	
S. cerevisiae AroA - Z48179.1 pro	(271)	LRNLLNFGHSIGHAYEAILTPQALHGECSVSIGMVKEAELSRYFGILSPTQVARLS	-----	-----	-----	-----	-----	
S. pombe AroA - AL157734 pro	(263)	LRNLLNFGHSIGHAYEAILYPQILHGECAVAGMVKEAELARYLGILKPNAVGRLT	-----	-----	-----	-----	-----	
B.petasis AroA - BPEAROA pro	(1)	-----	-----	-----	-----	-----	-----	
Aeromonas salmonicida AroA - A18838 pro	(1)	-----	-----	-----	-----	-----	-----	
Haemophilus influenzae AroA - HEAAAROAUR pro	(1)	-----	-----	-----	-----	-----	-----	
Haemophilus somnis AroA - HEA3P1C pro	(1)	-----	-----	-----	-----	-----	-----	
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(1)	-----	-----	-----	-----	-----	-----	
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(1)	-----	-----	-----	-----	-----	-----	
P. multocida aroA - PMAROA pro	(1)	-----	-----	-----	-----	-----	-----	
Vibrio cholerae AroA pro	(1)	-----	-----	-----	-----	-----	-----	
Y. enterocolitica AroA - YEPSERCAROpro	(1)	-----	-----	-----	-----	-----	-----	
Yersinia pestis AroA - YEPAROApro	(1)	-----	-----	-----	-----	-----	-----	
Klebsiella pneumoniae aroA - KPAROA pro	(1)	-----	-----	-----	-----	-----	-----	
S. typhi AroA - ST5E3PS pro	(1)	-----	-----	-----	-----	-----	-----	
S. typhimurium aroA - STYAROApm pro	(1)	-----	-----	-----	-----	-----	-----	
Salmonella gallinarum AroA - STYSERARO pro	(1)	-----	-----	-----	-----	-----	-----	
Shigella dysenteriae AroA - SDU82268 pro	(1)	-----	-----	-----	-----	-----	-----	
E. coli AroA - ECAROA pro	(1)	-----	-----	-----	-----	-----	-----	
Shigella sonnei AroA - AF101225 pro	(1)	-----	-----	-----	-----	-----	-----	
Consensus	(276)	-----	-----	-----	-----	-----	-----	

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

	Section 7				
	331	340	350	360	370
Oryza sativa AroA gene - AP002542 pro	(1)	-	-	-	-
Lolium rigidum AroA - AF349754 pro	(1)	-	-	-	-
Z. mays AroA - ZMEPSPSpro	(1)	-	-	-	-
N. tabacum AroA partial - M61905 pro	(1)	-	-	-	-
Petunia hybrida AroA - PETAROA pro	(1)	MAQINNMAQGIQTTLNPNNSN -	PHKPQVPKSSFLVFGSK -	KLKNANSMLVLIK	KD
Tomato AroA - TOMAROA pro	(1)	MAQISSMAQGIQTLSLNS	SKTQKGPLVNSLFFF	SAKS LGVFKKK	
Arabidopsis thaliana AroA cDNA - AF360224 pro	(1)	MASSLTSKSILGCTKPA	ASSFLPSELRLLS	HSQTRKNFRQSWGLKK	S
Arabidopsis thaliana AroA gene ALEPSPS	(1)	MAQVSRICNGVQNP -	SLISLNLSKSSQRKSPLSVSLKTQ	QHPRAY PISSSWGLKK	S
B. napus AroA - X51475 pro	(1)	MAQSSRICHGVQNP	CVIISNLSKSNQNKS	PSVSLKTHOPR ---	ASSWGLKK
Agrobacterium CP4 partial AroA sequence	(1)	-	-	-	-
D.nodosus (VCS1001) aroA - DNEPS3PS pro	(1)	-	-	-	-
Corynebacterium glutamicum AroA - AF114233 pro	(2)	RKQSPTRRFQKHL	FPGANDRNLRSKVVRSRSSLR	TTDRSSAISVASI	SE SP
Pyrococcus abyssi AroA - CNSPAX02 pro	(1)	-	-	-	-
S. cerevisiae AroA - Z48179.1 pro	(326)	K1LVAYGLFVSPDEKWP	KELTLHKKTPLDILLKKM	SIDKKNEGSKKVVVILE	SIG
S. pombe AroA - AL157734 pro	(318)	KCLVSYNLFISVNDPKV	KVKKYASF	KHCPVEKLIEYMAVDKKNQGS	KKRIVILKAIG
B. pertussis AroA - BPEAROA pro	(1)	-	-	-	-
Aeromonas salmonicida AroA - A18838 pro	(1)	-	-	-	-
Haemophilus influenzae AroA - HEAAROAUR pro	(1)	-	-	-	-
Haemophilus somnis AroA - HEA3P1C pro	(1)	-	-	-	-
Pasteurella haemolytica NADC-D80 AroA - PHU03068 pro	(1)	-	-	-	-
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(1)	-	-	-	-
P. multocida aroA - PMAROA pro	(1)	-	-	-	-
Vibrio cholerae AroA pro	(1)	-	-	-	-
Y. enterocolitica AroA - YEPSERCAROpro	(1)	-	-	-	-
Yersinia pestis AroA - YEPAROApro	(1)	-	-	-	-
Klebsiella pneumoniae aroA - KPAROA pro	(1)	-	-	-	-
S. typhi AroA - ST5E3PS pro	(1)	-	-	-	-
S. typhimurium aroA - STYAROAPM pro	(1)	-	-	-	-
Salmonella gallinarum AroA - STYSERARO pro	(1)	-	-	-	-
Shigella dysenteriae AroA - SDU82268 pro	(1)	-	-	-	-
E. coli AroA - ECAROA pro	(1)	-	-	-	-
Shigella sonnei AroA - AF101225 pro	(1)	-	-	-	-
Consensus	(331)	-	-	-	-

FIGURE 3
Page 7

EPSP Synthase CDS protein alignment

	(386)	386	400	410	420	430	440	Section 8
Oryza sativa AroA gene - AP002542 pro	(1)	-						
Lolium rigidum AroA - AF349754 pro	(1)	-						
Z. mays AroA - ZMEPSPSPro	(1)	-						
N. tabacum AroA partial - M61905 pro	(1)	-						
Petunia hybrida AroA - PETAROA pro	(53)	S I FMQKFCS -- FRISASVATAQKPSE	VLQPIKE	SGT	KLPGSKS	KLPGSKS	SNRILL	
Tomato AroA - TOMAROA pro	(56)	SVLRVVRKSS - FRISASVATAEKPHE	VLXPIXD	SGT	KLPGSKS	KLPGSKS	SNRILL	
Arabidopsis thaliana AroA cDNA - AF360224 pro	(56)	DMLNGSEIRPVKVRASVSTA	EKASE	VLQPIRE	SGL	KLPGSKS	KLPGSKS	SNRILL
Arabidopsis thaliana AroA gene AtEPSPS	(55)	GMLTIGSELRLPKVMSVSTA	EKASE	VLQPIRE	SGL	KLPGSKS	KLPGSKS	SNRILL
B. napus AroA - X51475 pro	(51)	GTMNLNGSVIRPVKVTA	SVST	SEKASE	VLQPIRE	SGL	KLPGSKS	SNRILL
Agrobacterium CP4 partial AroA sequence	(1)	-						
Brucella melitensis biovar Abortus AroA - AF326475 pro	(14)	I SQSRGVSAPKCDCEKSM	SMHSACPKPAT	ARHSQAT	GE	TRPQDKS	STR	F
O. nodosus (MC1001) aroA - DNEPS3PS pro	(1)	-						
Corynebacterium glutamicum AroA - AF114233 pro	(57)	VVNKTNLPHITSIMV	FVSDSS	I	MMTNIWHTAPV	APGSD	TCGDKS	SHRALL
Pyrococcus abyssi AroA - CNSPAX02 pro	(1)	-						
S. cerevisiae AroA - Z48179_1 pro	(381)	KCYGDSAQFVSD	EDLRFIL	DET	VVFPFKDIP	DQOKV	I PPGSKS	SNRALL
S. pombe AroA - AL157734 pro	(373)	E TYEKHATVVSDD	DIRFILSRDV	KVDEFTRK	- SWDV	V TPPGSKS	SNRALL	
B. pertussis AroA - BPEAROA pro	(1)	-						
Aeromonas salmonicida AroA - A18838 pro	(1)	-						
Haemophilus influenzae AroA - HEAAROAUR pro	(1)	-						
Haemophilus somnus AroA - HEA3P1C pro	(1)	-						
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(1)	-						
Pasteurella haemolytica serotype 2 aroA - PHU08948 pro	(1)	-						
P. multocida aroA - PMAROA pro	(1)	-						
Vibrio cholerae AroA pro	(1)	-						
Y. enterocolitica AroA - YEPSERCARO pro	(1)	-						
Yersinia pestis AroA - YEPAROA pro	(1)	-						
Klebsiella pneumoniae aroA - KPAROA pro	(1)	-						
S. typhi AroA - ST5E3PS pro	(1)	-						
S. typhimurium aroA - STYAROAPM pro	(1)	-						
Salmonella gallinarum AroA - STYSERARO pro	(1)	-						
Shigella dysenteriae AroA - SDU82268 pro	(1)	-						
E. coli AroA - ECAROA pro	(1)	-						
Shigella sonnei AroA - AF101225 pro	(1)	-						
Consensus	(386)							

FIGURE 3
Page 8

EPSP Synthase CDS protein alignment

Section 9					
	441	450	460	470	480
Oryza sativa AroA gene - AP002542 pro	(441) -	-	-	-	-
Lolium rigidum AroA - AF349754 pro	(1) -	-	-	-	-
Z. mays AroA - ZMEPSPSpro	(35) ALLEGTTV DNLNSD	DLH	MLGALRTLG	SVEADK	VAKR
N. tabacum AroA partial- M61905 pro	(1) -	-	-	-	-
Petunia hybrida AroA - PETAROA pro	(106) ALLEGTTV DNLSSDD	H	MLGALKTLG	THVEE	DAQR
Tomato AroA - TOMAROA pro	(110) ALLEGRTV DNLSSDD	H	MLGALKTLG	THVEDDN	AVR
Arabidopsis thaliana AroA cDNA - AF360224 pro	(111) ALLEGTTV DNLNSDD	N	MLDALKLIG	ENR	RAVE
Arabidopsis thaliana AroA gene AEPSPS	(110) ALLEGTTV DNLNSDD	N	MLDALKLIG	ENR	RAVE
B. napus AroA - X51475 pro	(106) ALLEGTTV DNLNSDD	N	MLDALKLIG	ENR	RAVE
Agrobacterium CIP4 partial AroA sequence	(1) -	-	-	-	-
Brucella melitensis biovar Abortus AroA - AF326475 pro	(69) LASGKTR TGLLG	G	D	INTGR	Q
D. nodosus (VCS1001) aroA - DNEPS3PS pro	(35) ALAEGQTE	RGFLACAD	CLATRQLR	LGDIQREK	DVWV
Corynebacterium glutamicum AroA - AF114233 pro	(38) LLADSPK KMNPLIS	DTIASLD	MLTAKL	HELGATISW	EDGETV
Pneumococcus abyssi AroA - CNSPAX02 pro	(436) ALLEGQCK XNLLHSDDT	DLT	MLTAKL	HELGATISW	EDGETV
S. cerevisiae AroA - Z48179.1 pro	(426) ALANGTCT TNLHSDDTQFM	S	MLTAKL	HELGATISW	EDGETV
S. pombe AroA - AL157734 pro	(36) ALAEGSTE TGLLDSDDTRVMLA	LRQ	LGATPSWEDGET	VVKGNGG	-
B. pertussis AroA - BPEAROA pro	(33) ALAEGTTT TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Aeromonas salmonicida AroA - A18838 pro	(33) ALAEGTTK TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Haemophilus influenzae AroA - HE2AAROAUR pro	(33) ALAKGTTK TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Haemophilus somnis AroA - HEA3P1C pro	(33) ALAKGTTQ TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(33) ALATGTTQ TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
P. multocida AroA - PMAROA pro	(37) ALATGTTQ TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Vibrio cholerae AroA pro	(33) ALASGTTT TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Y. enterocolitica AroA - YEPSEERCARO pro	(34) ALAEGTTQ TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Yersinia pestis AroA - YEPAROA pro	(34) ALAEGTTQ TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Klebsiella pneumoniae AroA - KPAROA pro	(33) ALARGTTV TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
S. typhi AroA - ST5E3PS pro	(33) ALACGKTV TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
S. typhimurium AroA - STYAROAPM pro	(33) ALPCGKTA TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Salmonella gallinarum AroA - STYSERARO pro	(33) ALACGKTV TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Shigella dysenteriae AroA - SDU82268 pro	(33) ALAHGKTV TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
E. coli AroA - ECAROA pro	(33) ALAHGKTV TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Shigella sonnei AroA - AF101225 pro	(33) ALAHGKTV TNLLDSDDE	TRMLA	LTOLGK	YFQGEV	EDG
Consensus	(441) ALA G T ITNLLDSDDIRHML AL ALGV	AD	C V G	GG	GG

PCT/US02/21837

FIGURE 3
Page 9

WO 03/013226

PCT/US02/21837

CD8 protein alignment

FIGURE 3
Page 10

EPSP Synthase CDS protein alignment

	Section 11				
	551	560	570	580	590
Oryza sativa AroA gene - AP002542 pro	(551) L	OLGAD	CFLGTECPP	R K I GGLPGGKVK	SGS SSSQQLTALLMAP
Lolium rigidum AroA - AF349754 pro	(87) L	OLGAD	CFLGTDCCPP	R N I GGLPGGKVK	SGS SSSQQLTALLMAP
Z. mays AroA - ZMEPSPSpro	(79) L	OLGAD	CFLGTDCCPP	R N I GGLPGGKVK	SGS SSSQQLTALLMAP
N. tabacum AroA partial- M61905 pro	(140) L	OLGAD	CFLGTDCCPP	R N I GGLPGGKVK	SGS SSSQQLTALLMAP
Petunia hybrida AroA - PETAROA pro	(34) L	OLGAE	CFLGTKCPP	R VSKGGLPGGGKVK	SGS SSSQQLTALLMAP
Tomato AroA - TOMAROA pro	(212) L	OLGAE	CFLGTKCPP	R VSKGGLPGGGKVK	SGS SSSQQLTALLMAP
Arabidopsis thaliana AroA cDNA - AF360224 pro	(216) L	OLGAD	CSLGTNCCPP	R VSKGGLPGGGKVK	SGS SSSQQLTALLMAP
Arabidopsis thaliana AroA gene AEPSPS	(217) L	OLGAD	CTLGTNCCPP	R N NGGLPGGGKVK	SGS SSSQQLTALLMAP
B. napus AroA - X51475 pro	(216) L	OLGAD	CTLGTNCCPP	R N NGGLPGGGKVK	SGS SSSQQLTALLMAP
Agrobacterium CP4 partial AroA sequence	(212) L	OLGAD	CTLGTNCCPP	R N NGGLPGGGKVK	SGS SSSQQLTALLMAP
(45) -	(45) -	(45) -	(45) -	(45) -	(45) -
Brucella melitensis biovar Abortus AroA - AF326475 pro	(167) L	REMGVQV	AAEGERMPLTL	GPR - - T ENP T AY RVP	SSAQVK SAVTLAG
D.nodocus (VCS1001) aroA - DNEPSSPS pro	(133) L	VQMGAKIVSHSNFT	PLAHLISR	PLTGDYD	PSAQVK SAVTLAG
Conynebacterium glutamicum AroA - AF114233 pro	(206) L	RSLGVE	NN - - - - -	DGS SSSQQLTALLMAP	PLTGDYD
Pyrococcus abyssi AroA - CNSPAX02 pro	(128) L	RSLGVKV	ISGEK - - - - -	DGS SSSQQLTALLMAP	PLTGDYD
S. cerevisiae AroA - Z48179.1 pro	(542) L	RANGTKIYLN	EGSLP	NG - - - - -	DGS SSSQQLTALLMAP
S. pombe AroA - AL157734 pro	(530) L	RANGCEIN	YDQGS	YD SSSKNGLKG	DGS SSSQQLTALLMAP
B.peruensis AroA - BPEAROA pro	(133) L	RQFGAGI	YVGLGEAG	YVGLGEAG	DGS SSSQQLTALLMAP
Aeromonas salmonicida AroA - A18838 pro	(133) L	ALKGAHIIQY	LKK	PPR GEGSIRV	DGS SSSQQLTALLMAP
Haemophilus influenzae AroA - HEAAROAUR pro	(136) L	RQAGADIRY	LE	PPR DGPVR	DGS SSSQQLTALLMAP
Haemophilus somnis AroA - HEA3P1C pro	(136) L	RTQGANI	QY	PPR DGPVR	DGS SSSQQLTALLMAP
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(136) L	RQVGAEEQY	LE	PPR DGPVR	DGS SSSQQLTALLMAP
Pasteurella haemolytica serotype 2 aroA - PHU88948 pro	(136) L	RQVGAEEQY	LE	PPR DGPVR	DGS SSSQQLTALLMAP
P. multocida AroA - PMAROA pro	(143) L	CQAGAEI	QY	PPR DGPVR	DGS SSSQQLTALLMAP
Vibrio cholerae AroA pro	(133) L	RQGAQI	QY	PPR DGPVR	DGS SSSQQLTALLMAP
Y. enterocolitica AroA - YEPSERCAROpro	(134) L	RQGGAQI	QY	PPR DGPVR	DGS SSSQQLTALLMAP
Yersinia pestis AroA - YEPAROApro	(134) L	RQGGAQI	QY	PPR DGPVR	DGS SSSQQLTALLMAP
Klebsiella pneumoniae aroA - KPAROA pro	(133) L	RQGGAQI	QY	PPR DGPVR	DGS SSSQQLTALLMAP
S. typhi AroA - ST5E3PS pro	(133) L	RQGGAN	I	PPR DGPVR	DGS SSSQQLTALLMAP
S. typhimurium aroA - STYAROAOP pro	(133) L	RQGGAN	I	PPR DGPVR	DGS SSSQQLTALLMAP
Salmonella gallinarum AroA - STYSERARO pro	(133) L	RQGGAN	I	PPR DGPVR	DGS SSSQQLTALLMAP
Shigella dysenteriae AroA - SDU82268 pro	(133) L	RIGRAK	I	PPR DGPVR	DGS SSSQQLTALLMAP
E. coli AroA - ECAROA pro	(133) L	RLLGGAK	I	PPR DGPVR	DGS SSSQQLTALLMAP
Shigella sonnei AroA - AF101225 pro	(133) L	RLLGGAK	I	PPR DGPVR	DGS SSSQQLTALLMAP
Consensus	(551) L	RG Q GA	IDYLEQE	YPLRI	G G G V VDGS SSSQQLTALLMAP

FIGURE 3
Page 11

EPSP Synthase CDS protein alignment

		Section 12			
		620	640	660	680
Oryza sativa AroA gene - AP002542 pro	(605) LALGDVEIEIDKLI S I P Y E E T	TLRLM EFGVKA EHS DS	DRF Y KGG Q KY K XS PG		
Lolium rigidum AroA - AF349754 pro	(138) LALGDVEIEIDKLI S V P Y E E T	TLRLM EFGVKA EHS DS	DRF Y KGG Q K XS PG		
Z. mays AroA - ZMEPSPSP pro	(130) LALGDVEIEIDKLI S V P Y E E T	TLRLM EFGVKA EHS DS	DRF Y KGG Q KY K XS PG		
N. tabacum AroA partial- M61905 pro	(191) LALGDVEIEIDKLI S I P Y E E T	TLRLM EFGVKA EHS DS	DRF Y KGG Q KY K XS PG		
(85) LALGDVEIEIDKLI S V P Y E E T	TLKL M EFGVKA EHS DS	DRF Y KGG Q KY K XS PG			
Petunia hybrida AroA - PETAROA pro	(263) LALGDVEIEIDKLI S V P Y E E T	TLKL M EFGVKA EHS DS	DRF Y KGG Q KY K XS PG		
Tomato AroA - TOMAROA pro	(267) LALGDVEIEIDKLI S V P Y E E T	TLKL M EFGVVF EHS SG	DRF Y KGG Q KY K XS PG		
Arabidopsis thaliana AroA cDNA - AF360224 pro	(268) LALGDVEIEIDKLI S V P Y E E T	TLKL M EFGVSA EHS ES	DRF Y KGG Q KY K XS PG		
Arabidopsis thaliana AroA gene ATEPSRS	(267) LALGDVEIEIDKLI S V P Y E E T	TLKL M EFGVSV EHS DS	DRF Y KGG Q KY K XS PG		
B. napus AroA - X51475 pro	(263) LALGDVEIEIDKLI S V P Y E E T	TLKL M EFGVSA EHS DS	DRF Y KGG Q KY K XS PG		
Agrobacterium CP4 partial AroA sequence	(45) -				
Brucella melitensis biovar Abortus AroA - AF326475 pro	(215) -	LNTPG T T W PVMTRDHTK LQG F GAD	T VETD KDGVRH RIVGQG K L G Q		
D. nodosus (VCS1001) aroA - DNEPS3PS pro	(180) -	LLADG TTR H T C G S D	TRM L P L F G G A L E I K E -	Q	
Corynebacterium glutamicum AroA - AF114233 pro	(254) F KNGVTVK H G G R L S M P H I	DRSAG E E E E	SENQV VHPGEI L G R T W		
Pynococcus abyssi AroA - CNSPAX02 pro	(174) -	K G I L T V E L N P V S P Y I	T L K M E S F G E E F E R -	- - - - -	
S. cerevisiae AroA - Z481791 pro	(594) A E E P V T L L A L G G K P S K L Y D	D T K M E K F G I N V E T S T T E P Y T Y I P K G H Y I N P S			
S. pombe AroA - AL157734 pro	(582) A E Q P V T L K L A G G K P S Q L Y D	D T A M A S F G V N V T K S T T E E N T Y I P C G K Y N P P			
B. pertussis AroA - BPEAROA pro	(188) R E G Q D I T I E V G L S K P Y I	D T I T N L M A R F G V S V R R - D G R A F T I A R D A V Y R G P G			
Aeromonas salmonicida AroA - A18838 pro	(184) A P V I P R I H K G I V S K P Y I D I T L H M N S S G V V	E H - D N K L F Y K G N Q S I V S P G			
Haemophilus influenzae AroA - HEAAROAUR pro	(186) I A E N D T E I E I G I V S K P Y I D I T L A M R D F G V K V E N -	Q K F O V K G N Q S Y I S P N			
Haemophilus somnis AroA - HEA3P1C pro	(186) I E G D M E I E I G I V S K P Y I D I T P A M M K D F G I N V D -	I G K Q O Y Y I S P Q			
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(185) I A E G D M E I E I I G I V S K P Y I D I T L S M N D F G I T V E N -	I D N Q F E F I V G K Q G Y V A P Q			
Pasteurella haemolytica serotype 2 aroA - PHU089948 pro	(186) I A E S D M E I E I I G I V S K P Y I D I T L S M N D F G I T V E N -	R D K T F V K G K Q G Y V A P Q			
P. multocida AroA - PMAROA pro	(193) I S A E D T E I E I G I V S K P Y I D I T L M Q T F G V E V E N -	Q A Q R F E V K G H Q Q Y Q S P H			
Vibrio cholerae AroA pro	(183) I A Q Q K V T I K I G I V S K P Y I D I T L M Q E Q F G V Q V I N -	I H D Q F E F I V G Q S Y V S P G			
Y. enterocolitica AroA - YEPSEERCARO pro	(182) I A E Q D T E I Q I O G I V S K P Y I D I T L H L M K A F G V D V V H -	E N Q F H K G G Q T Y R S P G			
Yersinia pestis AroA - YEPAROA pro	(183) I A E Q D T T I R I G I V S K P Y I D I T L H L M K A F G V D V V H -	E N Q F H K G G Q T Y R S P G			
Klebsiella pneumoniae AroA - KPAROA pro	(182) I A P Q D T V I A I K G I L V S P Y I D I T L H L M K T F G V E V E N -	Q A Q R F E V K G N Q Q Y Q S P G			
S. typhi AroA - ST5E3PS pro	(182) I A P E D T I R I K G I L V S K P Y I D I T L N L M K T F G V E V E N -	A N - H H Q F V K G G Q Q Y H S P G			
S. typhimurium AroA - STYAROAPM pro	(182) I A P K D T I I R K G I L V S K P Y I D I T L N L M K T F G V E V E N -	A N - H H Q F V K G G Q Q Y H S P G			
Salmonella gallinarum AroA - STYSERARO pro	(182) I A P K D T I I R K G I L V S K P Y I D I T L N L M K T F G V E V E N -	A N - H H Q F V K G G Q Q Y H S P G			
Shigella dysenteriae AroA - SDU82268 pro	(182) I A P E D T V I R I K G I L V S K P Y I D I T L N L M K T F G V E V E N -	O H Q Q F V K G G Q S Y Q S P G			
E. coli AroA - ECAROA pro	(182) I A P E D T V I R I K G I L V S K P Y I D I T L N L M K T F G V E V E N -	O H Q Q F V K G G Q S Y Q S P G			
Shigella sonnei AroA - AF101225 pro	(182) I A P E D T V I R I K G I L V S K P Y I D I T L N L M K T F G V E V E N -	Q H Q Q F V K G G Q S Y Q S P G			
Consensus	(606) L A D I I G E L V S K P Y I D I T L M F G V V E Y F V V K G G Q Y S P G				

FIGURE 3
Page 12

EPSP Synthase CDS protein alignment

Section 13									
	661	670	680	690	700	710	720	730	740
<i>Oryza sativa</i> AroA gene - AP002542 pro	(193) N - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>Lolium rigidum</i> AroA - AF349754 pro	(185) N - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>Z. mays</i> AroA - ZMEPSPSPro	(246) N - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>N. tabacum</i> AroA partial - M61905 pro	(140) K - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>Petunia hybrida</i> AroA - PETAROA pro	(318) K - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>Tomato</i> AroA - TOMAROA pro	(322) K - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>Arabidopsis thaliana</i> AroA cDNA - AF360224 pro	(323) N - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>Arabidopsis thaliana</i> AroA gene AtEPSPS	(322) N - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
<i>B. napus</i> AroA - X51475 pro	(318) N - AYVEGDAASSASYFLAA	AA ITGGT	TVQGC	GTTSIQGD	DF	FAVLEM	-	-	- MG
Agrobacterium CP4 partial AroA sequence	(45) -	-	-	-	-	-	-	-	-
<i>Brucella melitensis</i> biovar <i>Abortus</i> AroA - AF326475 pro	(268) T IDPGDPSSAFAFLPLP	AA	EGSE	VTNLMNPTRG	ITLQEMGADIE	ID	-	-	-
<i>D. nodosus</i> (VCS1001) aroA - DNEPSP3PS pro	(228) VLDPGDLE	AA	PLP	AA	ITLQEMGADIE	ID	-	-	-
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(307) R - IEPDLM-A	PLP	AA	PLP	AA	ITLQEMGADIE	ID	-	- MG
<i>Pyrococcus abyssi</i> AroA - CN5PAX02 pro	(220) SKEHVP	PGD	SSAS	PLP	AA	ITLQEMGADIE	ID	-	- FG
<i>S. cerevisiae</i> AroA - Z48179.1 pro	(649) E -	ESD	ASSA	PLP	AA	ITLQEMGADIE	ID	-	- MG
<i>S. pombe</i> AroA - AL157734 pro	(637) H -	VE	ESD	ASSA	PLP	AA	ITLQEMGADIE	ID	- MG
<i>B. pertussis</i> AroA - BPEAROA pro	(242) R -	MAE	GDA	SSAS	YFL	AA	ITLQEMGADIE	ID	- MG
<i>Aeromonas salmonicida</i> AroA - A18838 pro	(238) D -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Haemophilus influenzae</i> AroA - HEA2004 pro	(240) K -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Haemophilus somnis</i> AroA - HEA3P1C pro	(240) T -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Pasteurella haemolytica</i> NADC-D60 AroA - PHU03068 pro	(240) GNTV	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Pasteurella haemolytica</i> serotype 2 aroA - PHU89948 pro	(240) GNTV	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>P. multocida</i> aroA - PMAROA pro	(247) R -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Vibrio cholerae</i> AroA pro	(237) Q -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Y. enterocolitica</i> AroA - YEPSERCARO pro	(236) I -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Yersinia pestis</i> AroA - YEPAROA pro	(237) T -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(236) D -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>S. typhi</i> AroA - ST5E3PS pro	(236) R -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>S. typhimurium</i> aroA - STYAROApm pro	(236) R -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(236) R -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Shigella dysenteriae</i> AroA - SDU8228 pro	(236) T -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>E. coli</i> AroA - ECAROA pro	(236) T -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
<i>Shigella sonnei</i> AroA - AF101225 pro	(236) T -	VEG	DASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG
Consensus	(661)	YLV	EGDASSA	SASYFL	AA	ITLQEMGADIE	ID	-	- MG

FIGURE 3
Page 13

EPSP Synthase CDS protein alignment

Section 14									
	716	718	720	722	724	726	728	730	732
Oryza sativa AroA gene - AP002542 pro	(242) A[K]TWT[T]S	TVTGP	PR	---	E PYG	KHLKAID	D	NMNKP	PDVAMT
Lolium rigidum AroA - AF349754 pro	(234) A[K]TWT[T]S	TVTGP	PR	---	Q PEG	KHLKAID	D	NMNKP	PDVAMT
Z. mays AroA - 2MEPSPS pro	(295) A[K]TWT[T]S	TVTGP	PR	---	E PEG	KHLKAID	D	NMNKP	PDVAMT
N. tabacum AroA partial-M61905 pro	(189) A[E]TWT[N]S	TVKGP	PR	---	---	NSSAMKHLRAID	D	NMNKP	PDVAMT
Petunia hybrida AroA - PETAROA pro	(367) A[E]TWT[N]S	TVKGP	PR	---	SSSG	KHLRAID	D	NMNKP	PDVAMT
Tomato AroA - TOMAROA pro	(371) A[E]TWT[N]S	TVTGP	PR	---	---	NSSGMKHLRAID	D	NMNKP	PDVAMT
Arabidopsis thaliana AroA cDNA - AF380224 pro	(372) C[K]A[W]T[N]S	TVTGP	SR	---	DAFGMRHLRAID	D	NMNKP	PDVAMT	AVVALFA[G]P
Arabidopsis thaliana AroA gene AEPSPS	(371) C[K]A[W]T[N]S	TVTGP	PR	---	DAFGMRHLRAID	D	NMNKP	PDVAMT	AVVALFA[G]P
B. napus AroA - X51475 pro	(367) C[K]A[W]T[N]S	TVTGP	SR	---	DAFGMRHLRAID	D	NMNKP	PDVAMT	AVVALFA[G]P
Agrobacterium CP4 partial AroA sequence	(45) ---	---	---	---	---	---	---	---	---
Brucella melitensis biovar Abortus AroA - AF326475 pro	(323) P[R]AGG[D]YADLVRKAS	---	---	---	---	---	---	---	---
D. nodosus (VCS1001) aroA - DNEPS3PS pro	(283) Q[R]FWG[R]P	VAADIVVYHS	---	---	---	---	---	---	---
Corynebacterium glutamicum AroA - AF114233 pro	(354) C[E]ELVA[Q]GEGYDLSVTG	---	---	---	---	---	---	---	---
Pyrococcus abyssi AroA - CNSPAX02 pro	(289) AD[KVGRK]VVE	---	---	---	---	---	---	---	---
S. cerevisiae AroA - Z48179.1 pro	(659) CK[TQATAT]TVSGPPVGTLKPLKHDVEPT	---	---	---	---	---	---	---	---
S. pombe AroA - AL157734 pro	(687) C[T]EQATATSTTVQGPPKGTLKPLKESIDE	---	---	---	---	---	---	---	---
B. pertussis AroA - BPEAROA pro	(291) AD[R]QPGK[RETRGV]V	---	---	---	---	---	---	---	---
Aeromonas salmonicida AroA - A18838 pro	(285) AR[TWG]D	EAR	---	---	---	---	---	---	---
Haemophilus influenzae AroA - HEAAROAUR pro	(288) AK[TWG]D	QAE	---	---	---	---	---	---	---
Haemophilus somnus AroA - HEA3PIC pro	(288) AK[TWG]D	QVE	---	---	---	---	---	---	---
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(288) AK[TWG]D	QAB	---	---	---	---	---	---	---
P. multocida aroA - PMAROA pro	(289) AK[TWG]D	QAB	---	---	---	---	---	---	---
Vibrio cholerae serotype 2 aroA - PHU89848 pro	(295) AH[TWG]D	QVE	---	---	---	---	---	---	---
P. multocida aroA - PMAROA pro	(286) AQ[BNG]D	IAR	---	---	---	---	---	---	---
Y. enterocolitica AroA - YEPSERCARO pro	(285) AR[TWG]D	ECS	---	---	---	---	---	---	---
Klebsiella pneumoniae aroA - KPAROA pro	(286) AX[TWG]D	ECS	---	---	---	---	---	---	---
S. typhi AroA - STSESPS pro	(285) AT[TWG]D	ACT	---	---	---	---	---	---	---
S. typhimurium aroA - STYAROA pro	(285) AT[TWG]D	ACT	---	---	---	---	---	---	---
Salmonella gallinarum AroA - STYSERARO pro	(285) AT[TWG]D	ACT	---	---	---	---	---	---	---
Shigella dysenteriae AroA - SDU82268 pro	(285) AT[CWG]D	SCT	---	---	---	---	---	---	---
E. coli AroA - ECAROA pro	(285) AT[CWG]D	SCT	---	---	---	---	---	---	---
Shigella sonnei AroA - AF101225 pro	(285) AT[CWG]D	SCT	---	---	---	---	---	---	---
Consensus	(716) A	VTWGEDFI	---	---	---	---	---	---	---

EPSP synthase CDS protein alignment

	Section 15			
	771	771	780	790
				800
Oryza sativa AroA gene - AP002542 pro	(294) T A I R D T A S W R V K E T R I V A I R T E L T K G A S V E E G P -			
Lolium rigidum AroA - AF349754 pro	(286) T A I R D T A S W R V K E T R I V A C T E L T K G A T V E E G P -			
Z. mays AroA - ZMEPSPS pro	(347) T A I R D T A S W R V K E T R I V A C T E L T K G A S V E E G P -			
N. tabacum AroA partial - MB1905 pro	(241) T A I R D T A S W R V K E T R I V A C T E L T K G A T V E E G P -			
Petunia hybrida AroA - PETAROA pro	(419) T A I R D T A S W R V K E T R I V A C T E L T K G A T V E E G P -			
Tomato AroA - TOMAROA pro	(423) T T I R D T A S W R V K E T R I V A C T E L T K G A T V V E G S -			
Arabidopsis thaliana AroA cDNA - AF360224 pro	(424) T T I R D T A S W R V K E T R I V A C T E L T K G A T V E E G S -			
Arabidopsis thaliana AroA gene ATEPSPS	(423) T T I R D T A S W R V K E T R I V A C T E L T K G A T V E E G S -			
B. napus AroA - X51475 pro	(419) T T I R D T A S W R V K E T R I V A C T E L T K G A T V E E G S -			
Agrobacterium CP4 partial AroA sequence	(45) -			
Brucella melitensis biovar Abortus AroA - AF326475 pro	(372) T V T D G D D E L R V K E T D R L A A T A R G L E A N G V D C T E G E M S -			
D. nodosus (VCS1001) aroA - DNEPSJPS pro	(332) T F T G N A T S E L R V K E T D R L A A M A Q N L O T T G V A C T V G A -			
Corynebacterium glutamicum AroA - AF114233 pro	(403) T R T G I A H L R G H E T D R L A A T T A E T N K G T K C T E L K -			
Pyrococcus abyssi AroA - CNSPAX02 pro	(312) K S V I T G R A L R T K E D R T K A V N L R K A G I K V K E L P -			
S. cerevisiae AroA - Z481778.1 pro	(754) T T I E G I A N Q R V K E C N R T L A M A T E L A K F G V K T T E L P D G I Q V H G L N S I K D T K V S E D S			
S. pombe AroA - AL157734 pro	(741) T R I T G I A N Q R V K E C N R T L A M A T E L A K F G V T T E L P D G I Q V H G L N S I K D T K V S E D S			
B. pertussis AroA - BPEAROA pro	(341) C R T R N I G S W R V K E T D R T H A M H T E L K G A G V Q S G A -			
Aeromonas salmonicida AroA - A18838 pro	(328) V P P H S O Q H L Q L L A V R D D R C T P C T H G H R A Q A G G S E G -			
Haemophilus influenzae AroA - HEA3P1C pro	(332) T V I R N I Y N W R V K E T D R L T A M A T E L R K G A E V E E G E -			
Haemophilus somnis AroA - HEA3P1C pro	(332) T V I R N I Y N W R V K E T D R L T A M A T E L R K G A E V E E G E -			
Pasteurella haemolytica NADC-D86 AroA - PHU03068 pro	(332) T V I R N I Y N W R V K E T D R L T A M A T E L R K G A E V E E G E -			
P. multocida AroA pro	(339) T V I R N I Y N W R V K E T D R L T A M A T E L R K G A E V E E G E -			
Vibrio cholerae AroA pro	(330) T A I R N T Y N W R V K E T D R L T A M A T E L R K G A E V E E G E -			
Y. enterocolitica AroA - YEP SERC ARO pro	(329) T V I R N I Y N W R V K E T D R L T A M A T E L R K G A E V E E G Q -			
Yersinia pestis AroA - YEPAROA pro	(330) T T I R N I Y N W R V K E T D R L T A M A T E L R K G A E V E E G E -			
Klebsiella pneumoniae AroA - KPAROA pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G E -			
S. typhi AroA - STSE3PS pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G H -			
S. typhimurium AroA - STYAROAPM pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G H -			
Salmonella gallinarum AroA - STYSERARO pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G H -			
Shigella dysenteriae AroA - SDU82268 pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G H -			
E. coli AroA - ECAROA pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G H -			
Shigella sonnei AroA - AF101225 pro	(329) T T I R N I Y N W R V K E T D R L F A M A T E L R K G A E V E E G H -			
Consensus	(771) T T I R N I Y N W R V K E T D R L A M A T E L R K G A E V E E G			

EPSP Synthase CDS protein alignment

					Section 16
					870
					860
Oryza sativa AroA gene - AP002542 pro	(826)	-L VTAIDTYDDH R M A M A F S L A A C D V P V T -			
Lolium rigidum AroA - AF349754 pro	(339)	-L VTAIDTYDDH R M A M A F S L A A C D V P V T -			
Z. mays AroA - ZMEPSPSpro	(331)	-L VTAIDTYDDH R M A M A F S L A A C D V P V T -			
N. tabacum AroA partial - M61905 pro	(392)	-L V T V I D T Y D D H R M A M A F S L A A C D V P V T -			
(286)	-L V T V I D T Y D D H R M A M A F S L A A C D V P V T -				
Petunia hybrida AroA pro	(464)	-L V T V I D T Y D D H R M A M A F S L A A C D V P V T -			
Tomato AroA - PETAROA pro	(468)	-L V T V I D T Y D D H R M A M A F S L A A C D V P V T -			
Arabidopsis thaliana AroA cDNA - AF356224 pro	(469)	-K P A I D T Y D D H R M A M A F S L A A C D V P V T -			
Arabidopsis thaliana AroA gene AIEPSPS	(468)	-K T A I D T Y D D H R M A M A F S L A A C D V P V T -			
B. napus AroA - X51475 pro	(464)	-K P A I D T Y D D H R M A M A F S L A A C D V P V T -			
Agrobacterium CP4 partial AroA sequence	(45)				
Brucella melitensis biovar Abortus AroA - AF326475 pro	(419)	G L G G E T T A T E L D H R M A M S F L A A C D V P V T -			
D. nodosus (VCS1001) aroA - DNEPS3PS pro	(378)	Q F L P A R A N E E G D H R M A M S L A A C D V A G E L L -			
Corynebacterium glutamicum AroA - AF114233 pro	(449)	-G A V W H E I A D H R M A T A G I L L A D V D G V Q -			
Pneumococcus abyssi AroA - CNSPAX02 pro	(357)	-R G F T E E E N D H R M A M M A M S L A A C D V D G V Q -			
S. cereviciae AroA - Z48179.1 pro	(809)	-S G P V G I C T Y D D H R M A M S F S L A A G M V N D E V A N P V R I C E R H C T G K T P G -			
S. pombe AroA - AL157734 pro	(794)	-G I Y T Y D D H R M A M S F S L A A C D V D G V Q -			
B. pertussis AroA - BPEAROA pro	(387)	G W R D A H I G T E D D H R M A M C F L A A F G P A A V R -			
Aeromonas salmonicida AroA - A18838 pro	(373)	P A P A R R D R H L Q A S R E A M C F S L A A L D I A V T -			
Haemophilus influenzae AroA - HEAAROAUR pro	(378)	Q F K H A N I E T Y N D H R M A M C F S L A A L N T P V T -			
Haemophilus somnis AroA - HEA3P1C pro	(378)	K F K H A I E T Y N D H R M A M C F S L A A L D T S V T -			
Pasteurella haemolytica NADC-D60 AroA - FHU03068 pro	(381)	N F K H A I E T Y N D H R M A M C F S L A A L N T E V T -			
P. multocida aroA - PMAROA pro	(379)	N F K H A I E T Y N D H R M A M C F S L A A L N T E V T -			
Vibrio cholerae AroA pro	(375)	-L I H A A I D T Y D D H R M A M C F S L A A L D T P V T -			
Y. enterocolitica AroA - YEPSERCAROpro	(374)	-L I H A A I G T Y N D H R M A M C F S L A A L D T P V T -			
Yersinia pestis AroA - YEPAROAp	(375)	-L I H A A I G T Y N D H R M A M C F S L A A L D T P V T -			
Klebsiella pneumoniae aroA - KPAROA pro	(374)	-L I H A A I G T Y N D H R M A M C F S L A A L D T P V T -			
S. typhi AroA - ST5E3PS pro	(374)	-L I H A A I G T Y N D H R M A M C F S L A A L D T P V T -			
S. typhimurium aroA - STYAROAPM pro	(374)	-L I H A A I G T Y N D H R M A M C F S L A A L D T P V T -			
Salmonella gallinarum AroA - STYSERARO pro	(374)	-L I H A A I G T Y N D H R M A M C F S L A A L D T P V T -			
Shigella dysenteriae AroA - SDU82268 pro	(374)	-I N F A I A T Y N D H R M A M C F S L A A L D T P V T -			
E. coli AroA - ECAROA pro	(374)	-I N F A I A T Y N D H R M A M C F S L A A L D T P V T -			
Shigella sonnei AroA - AF010225 pro	(374)	-I N F A I A T Y N D H R M A M C F S L A A L D T P V T -			
Consensus	(826)	I N A E I T Y D H R M A M C F S L A A L D T P V T -			

FIGURE 3
Page 16

EPSP Synthase CDS protein alignment

	881	890	900	910	920	935	Section 17
Oryza sativa AroA gene - AP002542 pro	(881) QVLSTEVVN						
Lolium rigidum AroA - AF349754 pro	(383) QVLSTEVVN						
Z. mays AroA - ZMEPSPS pro	(348) -						
N. tabacum AroA partial - M61905 pro	(436) QVLSTEVVN						
Petunia hybrida AroA - PETAROA pro	(330) QVLQY KH						
Tomato AroA - TOMAROA pro	(508) QVLQY KH						
Arabidopsis thaliana AroA cDNA - AF380224 pro	(512) QVLQY KH						
Arabidopsis thaliana AroA gene AIEPSPS	(513) QVLQY KH						
B. napus AroA - X51475 pro	(512) QVLQY KH						
Agrobacterium CP4 partial AroA sequence	(508) QVLQY KH						
Brucella melitensis biovar Abortus AroA - AF326475 pro	(45) -						
D. nodosus (VCS1001) aroA - DNEPS3 pro	(465) GMAAG GAKIA BSGAE						
Corynebacterium glutamicum AroA - AF114233 pro	(424) SPAAA GMNVGEKDANKNHD						
Pneumococcus abyssi AroA - CNSPAX02 pro	(401) LDLRL LNEG						
S. cerevisiae AroA - Z48179.1 pro	(491) NVWEEV NVG-						
S. pombe AroA - AL157734 pro	(863) QVLHSELGAKL DGAEPLECTSKK						
B. pertussis AroA - BPEAROA pro	(835) NSKKSVVIIGMRAAGKTTISKWCASALGY						
Aeromonas salmonicida AroA - A18838 pro	(432) QVLHQSFGVKL TGATSVASDPLKGTSKNASIILIGMRGAGKTTIGKIIAKQLNF						
Haemophilus influenzae AroA - HEAAROAUR pro	(418) QVAG LLAARD						
Haemophilus somnius AroA - HEA3PIC pro	(423) NEFE CLKNN						
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(423) S EFE KKNQ						
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(426) RDLE QVVR						
P. multocida aroA - PMAROA pro	(424) R ELE QVVR						
Vibrio cholerae AroA pro	(429) I LFTLN REVAYR						
Y. enterocolitica AroA - YEPAROA pro	(418) QVAG QIA						
Yersinia pestis AroA - YEPAROA pro	(419) QVAG QIA						
Klebsiella pneumoniae aroA - KPAROA pro	(418) QVAG QIA						
S. typhi AroA - ST5E3PS pro	(418) QVAG QIA						
S. typhimurium aroA - STYAROAPM pro	(418) QVAG QIA						
Salmonella gallinarum AroA - STYSERRARO pro	(418) QVAG QIA						
Shigella dysenteriae AroA - SDU82268 pro	(418) QVAG QIA						
E. coli AroA - ECAROA pro	(418) QVAG QIA						
Shigella sonnei AroA - AF101225 pro	(418) QVAG QIA						
Consensus	(881) D L RIS						

FIGURE 3
Page 17

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

		936	935	934	933	932	931	930	929	928	927	926	925	924	923	922	921	920	919	918	917	916	915	914	913	912	911	910	909	908	907	906	905	904	903	902	901	900	899	898	897	896	895	894	893	892	891	890	889	888	887	886	885	884	883	882	881	880	879	878	877	876	875	874	873	872	871	870	869	868	867	866	865	864	863	862	861	860	859	858	857	856	855	854	853	852	851	850	849	848	847	846	845	844	843	842	841	840	839	838	837	836	835	834	833	832	831	830	829	828	827	826	825	824	823	822	821	820	819	818	817	816	815	814	813	812	811	810	809	808	807	806	805	804	803	802	801	800	799	798	797	796	795	794	793	792	791	790	789	788	787	786	785	784	783	782	781	780	779	778	777	776	775	774	773	772	771	770	769	768	767	766	765	764	763	762	761	760	759	758	757	756	755	754	753	752	751	750	749	748	747	746	745	744	743	742	741	740	739	738	737	736	735	734	733	732	731	730	729	728	727	726	725	724	723	722	721	720	719	718	717	716	715	714	713	712	711	710	709	708	707	706	705	704	703	702	701	700	699	698	697	696	695	694	693	692	691	690	689	688	687	686	685	684	683	682	681	680	679	678	677	676	675	674	673	672	671	670	669	668	667	666	665	664	663	662	661	660	659	658	657	656	655	654	653	652	651	650	649	648	647	646	645	644	643	642	641	640	639	638	637	636	635	634	633	632	631	630	629	628	627	626	625	624	623	622	621	620	619	618	617	616	615	614	613	612	611	610	609	608	607	606	605	604	603	602	601	600	599	598	597	596	595	594	593	592	591	590	589	588	587	586	585	584	583	582	581	580	579	578	577	576	575	574	573	572	571	570	569	568	567	566	565	564	563	562	561	560	559	558	557	556	555	554	553	552	551	550	549	548	547	546	545	544	543	542	541	540	539	538	537	536	535	534	533	532	531	530	529	528	527	526	525	524	523	522	521	520	519	518	517	516	515	514	513	512	511	510	509	508	507	506	505	504	503	502	501	500	499	498	497	496	495	494	493	492	491	490	489	488	487	486	485	484	483	482	481	480	479	
--	--	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

	Section 19				
	991	1000	1010	1020	1030
Oryza sativa AroA gene - AP002542 pro	(991)	-----	-----	-----	1045
Lolium rigidum AroA - AF349734 pro	(392)	-----	-----	-----	-----
Z. mays AroA - ZMEPSPSPro	(348)	-----	-----	-----	-----
N. tabacum AroA partial - M61905 pro	(445)	-----	-----	-----	-----
Petunia hybrida AroA - PETAROA pro	(539)	-----	-----	-----	-----
Tomato AroA - TOMAROA pro	(517)	-----	-----	-----	-----
Arabidopsis thaliana AroA cDNA - AF360224 pro	(521)	-----	-----	-----	-----
Arabidopsis thaliana AroA gene AIEPSPS	(622)	-----	-----	-----	-----
B. napus AroA - X51475 pro	(521)	-----	-----	-----	-----
Agrobacterium CP4 partial AroA sequence	(517)	-----	-----	-----	-----
Bacillus melitensis biovar Abortus AroA - AF326475 pro	(45)	-----	-----	-----	-----
D.nodosus (VCS1001) aroA - DNEPSPSPS pro	(481)	-----	-----	-----	-----
Corynebacterium glutamicum AroA - AF114233 pro	(444)	-----	-----	-----	-----
Pyrococcus abyssi AroA - CNSPAX02 pro	(499)	-----	-----	-----	-----
S. cerevisiae AroA - Z248179.1 pro	(411)	-----	-----	-----	-----
S. pombe AroA - AL157734 pro	(970)	GGIVESABRSRKALKDFASSGGYVLLHHRDIEETIVFLOSDPSRP	-AYVEEIREV	-----	-----
B. pertussis AroA - BPEAROA pro	(943)	GGVIEDMERSNLLSNFVKEGGIVLHVRNLIEHKSYLSEDQTRPTYKDQESIDDV	-----	-----	-----
Aeromonas salmonicida AroA - A18838 pro	(443)	-----	-----	-----	-----
Haemophilus influenzae AroA - HEAAROAUR pro	(428)	-----	-----	-----	-----
Haemophilus somnius AroA - HEA3P1C pro	(433)	-----	-----	-----	-----
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435)	-----	-----	-----	-----
Pasteurella haemolytica serotype 2 aroA - PHU88948 pro	(433)	-----	-----	-----	-----
P. multocida aroA - PMAROA pro	(433)	-----	-----	-----	-----
Vibrio cholerae AroA pro	(442)	-----	-----	-----	-----
Y. enterocolitica AroA - YEPSEERCAROpro	(427)	-----	-----	-----	-----
Yersinia pestis AroA - YEPAROApro	(428)	-----	-----	-----	-----
Klebsiella pneumoniae aroA - KPAROA pro	(425)	-----	-----	-----	-----
S. typhi AroA - ST5E3PS pro	(428)	-----	-----	-----	-----
S. typhimurium aroA - STYAROAPM pro	(428)	-----	-----	-----	-----
Salmonella gallinarum AroA - STYSERARO pro	(428)	-----	-----	-----	-----
Shigella dysenteriae AroA - SDU82268 pro	(428)	-----	-----	-----	-----
E. coli AroA - ECAROA pro	(428)	-----	-----	-----	-----
Shigella sonnei AroA - AF101225 pro	(428)	-----	-----	-----	-----
Consensus	(991)	-----	-----	-----	-----

EPSP Synthase CDS protein alignment

		Section 20						
		1046	1046	1060	1070	1080	1090	1100
Oryza sativa AroA gene - AP002542 pro	(392)	-	-	-	-	-	-	-
Lolium rigidum AroA - AF349754 pro	(348)	-	-	-	-	-	-	-
Z. mays AroA - ZMEPSPSpro	(445)	-	-	-	-	-	-	-
N. tabacum AroA partial- M61905 pro	(339)	-	-	-	-	-	-	-
Petunia hybrida AroA - PETAROA pro	(517)	-	-	-	-	-	-	-
Tomato AroA - TOMAROA pro	(521)	-	-	-	-	-	-	-
Arabidopsis thaliana AroA cDNA - AF360224 pro	(522)	-	-	-	-	-	-	-
Arabidopsis thaliana AroA gene AIEPSPS	(521)	-	-	-	-	-	-	-
B. napus AroA - X51475 pro	(517)	-	-	-	-	-	-	-
Agrobacterium CP4 partial AroA sequence	(45)	-	-	-	-	-	-	-
Brucella melitensis biovar Abortus AroA - AF326475 pro	(481)	-	-	-	-	-	-	-
D.nodosus (VCS1001) aroA - DNEPSPS3PS pro	(444)	-	-	-	-	-	-	-
Corynebacterium glutamicum AroA - AF114233 pro	(499)	-	-	-	-	-	-	-
Pyrococcus abyssi AroA - CNSPAX02 pro	(411)	-	-	-	-	-	-	-
S. cerevisiae AroA - Z48178.1 pro	(1023)	NNRREGWYKECSNFSFPAPHCSAEEFQALRRSPSKYIATITGVREIEIPSGR-	-	-	-	-	-	-
S. pombe AroA - AL157734 pro	(998)	YKRRHVVWYRECRSHYFISPVLSNOVIDEKIQYSMSRFLDVVTGSSQVLQFKTKK	-	-	-	-	-	-
B. pertussis AroA - BPEAROA pro	(443)	-	-	-	-	-	-	-
Aeromonas salmonicida AroA - A18838 pro	(428)	-	-	-	-	-	-	-
Haemophilus influenzae AroA - HEA20AUR pro	(433)	-	-	-	-	-	-	-
Haemophilus somnis AroA - HEA3PIC pro	(433)	-	-	-	-	-	-	-
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435)	-	-	-	-	-	-	-
Pasteurella haemolytica serotype 2 aroA - PHU89984 pro	(433)	-	-	-	-	-	-	-
P. multocida aroA - PMAROA pro	(442)	-	-	-	-	-	-	-
Vibrio cholerae AroA pro	(427)	-	-	-	-	-	-	-
Y. enterocolitica AroA - YEPSERCAROpro	(428)	-	-	-	-	-	-	-
Yersinia pestis AroA - YEPAROA pro	(425)	-	-	-	-	-	-	-
Klebsiella pneumoniae aroA - KPAROA pro	(428)	-	-	-	-	-	-	-
S. typhi AroA - ST5E3PS pro	(428)	-	-	-	-	-	-	-
S. typhimurium aroA - STYAROAPM pro	(428)	-	-	-	-	-	-	-
Salmonella gallinarum AroA - STYSERARO pro	(428)	-	-	-	-	-	-	-
Shigella dysenteriae AroA - SDU82288 pro	(428)	-	-	-	-	-	-	-
E. coli AroA - ECAROA pro	(428)	-	-	-	-	-	-	-
Shigella sonnei AroA - AF101225 pro	(428)	-	-	-	-	-	-	-
Consensus (1046)								

WO 03/013226

PCT/US02/21837

EPSP synthase CDS protein alignment

	Section 21				
	1101	1110	1120	1130	1140
Oryza sativa AroA gene - AP002542 pro	(1101) -	(392) -			
Lolium rigidum AroA - AF349754 pro	(348) -				
Z. mays AroA - ZMEPSPSpro	(445) -				
N. tabacum AroA partial - M61905 pro	(339) -				
Petunia hybrida AroA - PETAROA pro	(517) -				
Tomato AroA - TOMAROA pro	(521) -				
Arabidopsis thaliana AroA cDNA - AF360224 pro	(522) -				
Arabidopsis thaliana AroA gene AEPSPS	(521) -				
B. napus AroA - X51475 pro	(517) -				
Agrobacterium CP4 partial AroA sequence	(45) -				
Brucella melitensis biovar Abortus AroA - AF326475 pro	(481) -				
D.nodosus (VCS1001) aroA - DNEPSP3PS pro	(444) -				
Corynebacterium glutamicum AroA - AF114233 pro	(499) -				
Pyrococcus abyssi AroA - CNSPAX02 pro	(411) -				
S. cerevisiae AroA - Z48179.1 pro	(1076) -	SAPVCLTFDDLT EQTENLTPI CYGCEA E V E V R V D H L A N -	-	-	-
S. pombe AroA - AL157734 pro	(1053) -	RSTFLTLNYPRIEDALPTLRDVYLGDAIE V R V D Y L K D P K S S N G I S S L D F V A E Q I	-	-	-
B. pertussis AroA - BPEAROA pro	(443) -				
Aeromonas salmonicida AroA - A18838 pro	(428) -				
Haemophilus influenzae AroA - HEARAOAUR pro	(433) -				
Haemophilus somnis AroA - HEA3P1C pro	(433) -				
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435) -				
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(433) -				
Vibrio cholerae AroA pro	(427) -				
P. multocida aroA - PMAROA pro	(442) -				
Y. enterocolitica AroA - YEP SERC ARO pro	(428) -				
Yersinia pestis AroA - YEPAROA pro	(425) -				
Klebsiella pneumoniae aroA - KPAROA pro	(428) -				
S. typhi AroA - ST5E3PS pro	(428) -				
S. typhimurium aroA - STYAROA PM pro	(428) -				
Salmonella gallinarum AroA - STY SERARAO pro	(428) -				
Shigella dysenteriae AroA - SDU82288 pro	(428) -				
E. coli AroA - ECAROA pro	(428) -				
Shigella sonnei AroA - AF101225 pro	(428) -				
Consensus	(1101)				

EPSP Synthase CDS protein alignment

		Section 22				
		1158	1170	1180	1180	1200
<i>Oryza sativa</i> AroA gene - AP002542 pro	(1156)	---	(392)	---	---	---
<i>Lolium rigidum</i> AroA - AF349754 pro	(348)	---	---	---	---	---
<i>Z. mays</i> AroA - ZMEPSPS pro	(445)	---	---	---	---	---
<i>N. tabacum</i> AroA partial M61905 pro	(339)	---	---	---	---	---
<i>Petunia hybrida</i> AroA - PETAROA pro	(517)	---	---	---	---	---
<i>Tomato</i> AroA - TOMAROA pro	(521)	---	---	---	---	---
<i>Arabidopsis thaliana</i> AroA cDNA - AF360224 pro	(522)	---	---	---	---	---
<i>Arabidopsis thaliana</i> AroA gene AIEPSPS	(521)	---	---	---	---	---
<i>B. napus</i> AroA - X51475 pro	(517)	---	---	---	---	---
<i>Agrobacterium</i> CP4 partial AroA sequence	(45)	---	---	---	---	---
<i>Bacillus</i> melitensis biovar Abortus AroA - AF326475 pro	(481)	---	---	---	---	---
<i>D. nodosus</i> (VCS1001) aroA - DNEPSSAPS pro	(444)	---	---	---	---	---
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(499)	---	---	---	---	---
<i>Pyrococcus abyssi</i> AroA - CNSPAJ02 pro	(411)	---	---	---	---	---
<i>S. cerevisiae</i> AroA - Z487179.1 pro	(1123)	---	---	---	---	---
<i>S. pombe</i> AroA - AL157734 pro	(1108)	---	---	---	---	---
<i>B. pertussis</i> AroA - BPEAROA pro	(443)	---	---	---	---	---
<i>Aeromonas salmonicida</i> AroA - A18838 pro	(428)	---	---	---	---	---
<i>Haemophilus influenzae</i> AroA - HEAAROAUR pro	(433)	---	---	---	---	---
<i>Haemophilus somnus</i> AroA - HEA3P1C pro	(433)	---	---	---	---	---
<i>Pasteuria haemolytica</i> NADC-D60 AroA - PHU03068 pro	(435)	---	---	---	---	---
<i>P. multocida</i> aroA - PMAROA pro	(433)	---	---	---	---	---
<i>Vibrio cholerae</i> AroA pro	(442)	---	---	---	---	---
<i>Y. enterocolitica</i> AroA - YEP SERC ARO pro	(428)	---	---	---	---	---
<i>Yersinia pestis</i> AroA - YEPAROA pro	(425)	---	---	---	---	---
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(428)	---	---	---	---	---
<i>S. typhi</i> AroA - ST5E3PS pro	(428)	---	---	---	---	---
<i>S. lyphimurium</i> aroA - STYAROAPM pro	(428)	---	---	---	---	---
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(428)	---	---	---	---	---
<i>Shigella dysenteriae</i> AroA - SDU82268 pro	(428)	---	---	---	---	---
<i>E. coli</i> AroA - ECAROA pro	(428)	---	---	---	---	---
<i>Shigella sonnei</i> AroA - AF101225 pro	(428)	---	---	---	---	---
Consensus (1156)						

FIGURE 3
Page 22

WO 03/013226

PCT/US02/21837

EPSP synthase CDS protein alignment

	Section 23						
	(1211)	1211	1220	1230	1240	1250	1265
Oryza sativa AroA gene - AP002542 pro	(392)	-	-	-	-	-	-
Lolium rigidum AroA - AF249754 pro	(348)	-	-	-	-	-	-
Z. mays AroA - ZMEPSPSPro	(445)	-	-	-	-	-	-
N. tabacum AroA partial - M61905 pro	(339)	-	-	-	-	-	-
Petunia hybrida AroA - PETAROA pro	(517)	-	-	-	-	-	-
Tomato AroA - TOMAROA pro	(521)	-	-	-	-	-	-
Arabidopsis thaliana AroA cDNA - AF360224 pro	(522)	-	-	-	-	-	-
Arabidopsis thaliana AroA gene AtEPSPS	(521)	-	-	-	-	-	-
B. napus AroA - X51475 pro	(517)	-	-	-	-	-	-
Agrobacterium CP4 partial AroA sequence	(45)	-	-	-	-	-	-
Brucella melitensis biovar Abortus AroA - AF326475 pro	(481)	-	-	-	-	-	-
D.nodosus (VCS1001) aroA - DNEPSS3PS pro	(444)	-	-	-	-	-	-
Corynebacterium glutamicum AroA - AF114233 pro	(499)	-	-	-	-	-	-
Pyrococcus abyssi AroA - CNSPAX02 pro	(411)	-	-	-	-	-	-
S. cerevisiae AroA - Z48179.1 pro	(1178)	TDIQYEVINRKGNNTKIGSHDFQGLYSWDDAEWENRFNQALTLDDVVKFVGTA	-	-	-	-	-
S. pombe AroA - AL157734 pro	(1162)	SETINILYQHKGYTKLIMSNWHDLSGTWSWARPHEWMQKVELASSYADVIKLVGMA	-	-	-	-	-
B.pertussis AroA - BPEAROA pro	(443)	-	-	-	-	-	-
Aeromonas salmonicida AroA - A18838 pro	(428)	-	-	-	-	-	-
Haemophilus influenzae AroA - HEAAROAUR pro	(433)	-	-	-	-	-	-
Haemophilus somnius AroA - HEA3P1C pro	(433)	-	-	-	-	-	-
Pasteurella haemolytica NADC-D80 AroA - PHU03068 pro	(435)	-	-	-	-	-	-
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(433)	-	-	-	-	-	-
P. multocida aroA - PMAROA pro	(442)	-	-	-	-	-	-
Vibrio cholerae AroA pro	(427)	-	-	-	-	-	-
Y. enterocolitica AroA - YEP SERC ARO pro	(428)	-	-	-	-	-	-
Yersinia pestis AroA - YEPAROA pro	(425)	-	-	-	-	-	-
Klebsiella pneumoniae aroA - KPAROA pro	(428)	-	-	-	-	-	-
S. typhi AroA - ST5E3PS pro	(428)	-	-	-	-	-	-
S. typhimurium aroA - STYAROAPM pro	(428)	-	-	-	-	-	-
Salmonella gallinarum AroA - STYSERARO pro	(428)	-	-	-	-	-	-
Shigella dysenteriae AroA - SDU82268 pro	(428)	-	-	-	-	-	-
E. coli AroA - ECAROA pro	(428)	-	-	-	-	-	-
Shigella sonnei AroA - AF101225 pro	(428)	-	-	-	-	-	-
Consensus (1211)							

FIGURE 3
Page 23

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

							Section 24
	(1266) 1266						
		(392) -					
<i>Oryza sativa</i> AroA gene - AP002542pro							
<i>Lolium rigidum</i> AroA - AF349754 pro	(348) -						
<i>Z. mays</i> AroA - ZMEPSPSpro	(445) -						
<i>N. tabacum</i> AroA partial - M61905 pro	(339) -						
<i>Petunia hybrida</i> AroA - PETAROA pro	(517) -						
<i>Tomato</i> AroA - TOMAROA pro	(521) -						
<i>Arabidopsis thaliana</i> AroA cDNA - AF360224 pro	(522) -						
<i>Arabidopsis thaliana</i> AroA gene ATEPSPS	(521) -						
<i>B. napus</i> AroA - X51475 pro	(517) -						
<i>Agrobacterium</i> CP4 partial AroA sequence	(45) -						
<i>Bacillus</i> mellitensis biovar Abortus AroA - AF326475 pro	(481) -						
<i>D. nodosus</i> (VCS1001) aroA - DNEPSSPS pro	(444) -						
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(499) -						
<i>Pyrococcus abyssi</i> AroA - CNNSPAX02 pro	(411) -						
<i>S. cerevisiae</i> AroA - Z48179.1 pro	(1233) -	VNFEDNLRLEHFRDTHKN --	KPLIAVNMTSKGSISRVVNNVLT	PVTSDLLPNSA			
<i>S. pombe</i> AroA - AL157734 pro	(1217) -	NNLNDNLEEEFRTRITNSMDIPLILFNMGRFGQLSRILNKFMTR	PVTTHPLLPNSKA				
<i>B. pertussis</i> AroA - BPEAROA pro	(443) -						
<i>Aeromonas salmonicida</i> AroA - A18838 pro	(428) -						
<i>Haemophilus influenzae</i> AroA - HEAAROAU pro	(433) -						
<i>Haemophilus somnius</i> AroA - HEA3P1C pro	(433) -						
<i>Pasteurella haemolytica</i> NADC-D60 AroA - PHU03068 pro	(435) -						
<i>Pasteurella haemolytica</i> serotype 2 aroA - PHU89948 pro	(433) -						
<i>P. multocida</i> aroA - PMAROA pro	(442) -						
<i>Vibrio cholerae</i> AroA pro	(427) -						
<i>Y. enterocolitica</i> AroA - YEPSERCAROpro	(428) -						
<i>Yersinia pestis</i> AroA - YEPAROApro	(425) -						
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(428) -						
<i>S. typhi</i> AroA - ST5E3PS pro	(428) -						
<i>S. typhimurium</i> aroA - STYAROAPM pro	(428) -						
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(428) -						
<i>Shigella dysenteriae</i> AroA - SDU82268 pro	(428) -						
<i>E. coli</i> AroA - ECAROA pro	(428) -						
<i>Shigella sonnei</i> AroA - AF101225 pro	(428) -						
	Consensus (1266)						

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

	Section 25						
	(1321)	1321	1330	1340	1350	1360	1375
Oryza sativa AroA gene - AP002542 pro	(392)	-					
Lolium rigidum AroA - AF349754 pro	(348)	-					
Z. mays AroA - ZMEPSPSpro	(445)	-					
N. tabacum AroA partial - M61905 pro	(339)	-					
Petunia hybrida AroA - PETAROA pro	(517)	-					
Tomato AroA - TOMAROA pro	(521)	-					
Arabidopsis thaliana AroA cDNA - AF360224 pro	(522)	-					
Arabidopsis thaliana AroA gene AtEPSPS	(521)	-					
B. napus AroA - X51475 pro	(517)	-					
Agrobacterium CP4 partial AroA sequence	(45)	-					
Brucella melitensis biovar Abortus AroA - AF326475 pro	(481)	-					
D. nodosus (VCS1001) aroA - DNEPS3PS pro	(444)	-					
Corynebacterium glutamicum AroA - AF114233 pro	(499)	-					
S. cerevisiae AroA - CNSPAX02 pro	(411)	-					
S. cerevisiae AroA - AL157734 pro	(1285)	APGQLTVAQINKMYSMGGIEPKELFVVGKPIGHRSRSPILHNTGYEILGLPHKFD					
B. pombe AroA - BPEAROA pro	(1272)	APGQLTVKQNLNEARVLTGEILPEKFFLFGKPIKHSRSPILHSTAYELLGLPHTYE					
B. pertussis AroA - BPEAROA pro	(443)	-					
Aeromonas salmonicida AroA - A18838 pro	(428)	-					
Haemophilus influenzae AroA - HEAAROAUR pro	(433)	-					
Haemophilus somnius AroA - HEA3P1C pro	(433)	-					
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435)	-					
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(433)	-					
P. multocida aroA - PMAROA pro	(442)	-					
Vibrio cholerae AroA pro	(427)	-					
Y. enterocolitica AroA - YEPSERCARO pro	(428)	-					
Yersinia pestis AroA - YEPAROA pro	(425)	-					
Klebsiella pneumoniae aroA - KPAROA pro	(428)	-					
S. typhi AroA - ST5E3PS pro	(428)	-					
S. typhimurium aroA - STYAROAPM pro	(428)	-					
Salmonella gallinarum AroA - STYSERARO pro	(428)	-					
Shigella dysenteriae AroA - SDU82268 pro	(428)	-					
E. coli AroA - ECAROA pro	(428)	-					
Shigella sonnei AroA - AF101225 pro	(428)	-					
Consensus (1321)							

FIGURE 3
Page 25

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

								Section 26	
		(1376)	1376		1390	1400	1410	1420	1430
Oryza sativa AroA gene - AP002542 pro	(392)	-	-	-	-	-	-	-	-
Lolium rigidum AroA - AF349754 pro	(348)	-	-	-	-	-	-	-	-
Z. mays AroA - ZMEPSPSpro	(445)	-	-	-	-	-	-	-	-
N. tabacum AroA partial - M81905 pro	(339)	-	-	-	-	-	-	-	-
Petunia hybrida AroA - PETAROA pro	(517)	-	-	-	-	-	-	-	-
Tomato AroA - TOMAROA pro	(521)	-	-	-	-	-	-	-	-
Arabidopsis thaliana AroA cDNA - AF360224 pro	(522)	-	-	-	-	-	-	-	-
Arabidopsis thaliana AroA gene AIEPSPS	(521)	-	-	-	-	-	-	-	-
B. napus AroA - X51475 pro	(517)	-	-	-	-	-	-	-	-
Agrobacterium CP4 partial AroA sequence	(45)	-	-	-	-	-	-	-	-
Brucella melitensis biovar Abortus AroA - AF326475 pro	(481)	-	-	-	-	-	-	-	-
D. nodosus (VCS1001) aroA - DNEEPS3PS pro	(444)	-	-	-	-	-	-	-	-
Corynebacterium glutamicum AroA - AF114233 pro	(499)	-	-	-	-	-	-	-	-
Pyrococcus abyssi AroA - CNSPAX02 pro	(411)	-	-	-	-	-	-	-	-
S. cerevisiae AroA - Z48179.1 pro (1340)	(1340)	KFETESAQLVKEKLLDGNKNFGGAATTIPLKLDDIMQYMDDELTDAAKVIGAVNTVI							
S. pombe AroA - AL157734 pro (1327)	(1327)	AFETDTVDBVQKVNLNP--DFFGANVTIPYKL SVMKFMDELSDEARFFGAVNTII							
B. pertussis AroA - BPEAROA pro	(443)	-	-	-	-	-	-	-	-
Aeromonas salmonicida AroA - A18838 pro	(428)	-	-	-	-	-	-	-	-
Haemophilus influenzae AroA - HEAAROAU pro	(433)	-	-	-	-	-	-	-	-
Haemophilus somnius AroA - HEA3P1C pro	(433)	-	-	-	-	-	-	-	-
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435)	-	-	-	-	-	-	-	-
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(433)	-	-	-	-	-	-	-	-
P. multocida aroA - PMAROA pro	(442)	-	-	-	-	-	-	-	-
Vibrio cholerae AroA pro	(427)	-	-	-	-	-	-	-	-
Y. enterocolitica AroA - YEPSERCAROpro	(428)	-	-	-	-	-	-	-	-
Yersinia pestis AroA - YEPAROA pro	(425)	-	-	-	-	-	-	-	-
Klebsiella pneumoniae aroA - KPAROA pro	(428)	-	-	-	-	-	-	-	-
S. typhi AroA - ST5E5PS pro	(428)	-	-	-	-	-	-	-	-
S. typhimurium aroA - STYAROAPM pro	(428)	-	-	-	-	-	-	-	-
Salmonella gallinarum AroA - STYSERARO pro	(428)	-	-	-	-	-	-	-	-
Shigella dysenteriae AroA - SDU82268 pro	(428)	-	-	-	-	-	-	-	-
E. coli AroA - ECAROA pro	(428)	-	-	-	-	-	-	-	-
Shigella sonnei AroA - AF101225 pro	(428)	-	-	-	-	-	-	-	-
Consensus (1376)									

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

	1431	1440	1450	1460	1470	1485	Section 27
Oryza sativa AroA gene - AP002542 pro	(1431)	-----	-----	-----	-----	-----	
Lolium rigidum AroA - AF349754 pro	(392)	-----	-----	-----	-----	-----	
Z. mays AroA - ZMEPSPSPro	(348)	-----	-----	-----	-----	-----	
N. tabacum AroA partial - M61905 pro	(445)	-----	-----	-----	-----	-----	
Petunia hybrida AroA - PETAROA pro	(339)	-----	-----	-----	-----	-----	
Tomato AroA - TOMAROA pro	(517)	-----	-----	-----	-----	-----	
Arabidopsis thaliana AroA cDNA - AF360224 pro	(521)	-----	-----	-----	-----	-----	
Arabidopsis thaliana AroA gene A1EPSPS	(522)	-----	-----	-----	-----	-----	
B. napus AroA - X51475 pro	(521)	-----	-----	-----	-----	-----	
Agrobacterium CP4 partial AroA sequence	(45)	-----	-----	-----	-----	-----	
Bacillus melitensis biovar Abortus AroA - AF326475 pro	(481)	-----	-----	-----	-----	-----	
D.nodosus (VCS1001) aroA - DNEPSPS3PS pro	(444)	-----	-----	-----	-----	-----	
Corynebacterium glutamicum AroA - AF114233 pro	(499)	-----	-----	-----	-----	-----	
Pyrococcus abyssi AroA - CNSPAX02 pro	(411)	-----	-----	-----	-----	-----	
S. cerevisiae AroA - Z48179.1 pro	(1395)	PIGN-----	-----	-----	-----	-----	
S. pombe AroA - AL157734 pro	(1380)	P R I G D K L V L R G D N T D W R G I Y D T F A N A L D G V S L R D T N G L V I G A G G T S R A A I Y S L H	-----	-----	-----	-----	
B. pertussis AroA - BPEAROA pro	(443)	-----	-----	-----	-----	-----	
Aeromonas salmonicida AroA - A18638 pro	(428)	-----	-----	-----	-----	-----	
Haemophilus influenzae AroA - HEAAROAUR pro	(433)	-----	-----	-----	-----	-----	
Haemophilus somnius AroA - HEA3P1C pro	(433)	-----	-----	-----	-----	-----	
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435)	-----	-----	-----	-----	-----	
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(433)	-----	-----	-----	-----	-----	
P. multocida aroA - PMAROA pro	(442)	-----	-----	-----	-----	-----	
Vibrio cholerae AroA pro	(427)	-----	-----	-----	-----	-----	
Y. enterocolitica AroA - YEPSEERCAROpro	(428)	-----	-----	-----	-----	-----	
Yersinia pestis AroA - YEPAROApro	(425)	-----	-----	-----	-----	-----	
Klebsiella pneumoniae aroA - KPAROA pro	(428)	-----	-----	-----	-----	-----	
S. typhi AroA - ST5E3PS pro	(428)	-----	-----	-----	-----	-----	
S. typhimurium aroA - STY/AROAPM pro	(428)	-----	-----	-----	-----	-----	
Salmonella gallinarum AroA - STYSERARO pro	(428)	-----	-----	-----	-----	-----	
Shigella dysenteriae AroA - SDU82268 pro	(428)	-----	-----	-----	-----	-----	
E. coli AroA - ECAROA pro	(428)	-----	-----	-----	-----	-----	
Shigella sonnei AroA - AF101225 pro	(428)	-----	-----	-----	-----	-----	
Consensus (1431)							

EPSP Synthase CDS protein alignment

					Section 28					
					1486 1488	1500	1510	1520	1530	1540
<i>Oryza sativa</i> AroA gene - AP002542 pro	(392)	-								
<i>Lolium rigidum</i> AroA - AF349754 pro	(348)	-								
<i>Z. mays</i> AroA - ZMEPSPSpro	(445)	-								
<i>N. tabacum</i> AroA partial- M61905 pro	(339)	-								
<i>Petunia hybrida</i> AroA - PETAROA pro	(517)	-								
<i>Tomato</i> AroA - TOMAROA pro	(521)	-								
<i>Arabidopsis thaliana</i> AroA cDNA - AF380224 pro	(522)	-								
<i>Arabidopsis thaliana</i> AroA gene AIEPSPS	(521)	-								
<i>B. napus</i> AroA - X51475 pro	(517)	-								
<i>Agrobacterium</i> CP4 partial AroA sequence	(45)	-								
<i>Brucella melitensis</i> biovar Abortus AroA - AF326475 pro	(481)	-								
<i>D. nodosus</i> (VCS1001) aroA - DNEPS3PS pro	(444)	-								
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(499)	-								
<i>Pyrrococcus abyssi</i> AroA - CNSPAX02 pro	(411)	-								
<i>S. cerevisiae</i> AroA - Z48179.1 pro	(1447)	-	SLGCKKKIFIINRTTSKLKPLIESLPSEEFNIGIESTKS-	-IEEIKERHVGVAVSCV						
<i>S. pombe</i> AroA - AL157734 pro	(1435)	-	RLGVSRIVLLNRTLANSYRVQDVFPPBDYNHIIIDSNDNIPSEELSSVTLSAVVSTI							
<i>B. pertussis</i> AroA - BPEAROA pro	(443)	-								
<i>Aeromonas salmonicida</i> AroA - A18838 pro	(428)	-								
<i>Haemophilus influenzae</i> AroA - HEARAOAUR pro	(433)	-								
<i>Haemophilus somnius</i> AroA - HEA3P1C pro	(433)	-								
<i>Pasteurella haemolytica</i> NADC-D60 AroA - PHU03068 pro	(435)	-								
<i>Pasteurella haemolytica</i> serotype 2 aroA - PHU89948 pro	(433)	-								
<i>P. multocida</i> aroA - PMAROA pro	(442)	-								
<i>Vibrio cholerae</i> AroA pro	(427)	-								
<i>Y. enterocolitica</i> AroA - YEPSERCARO pro	(428)	-								
<i>Yersinia pestis</i> AroA - YEPAROA pro	(425)	-								
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(428)	-								
<i>S. Typhi</i> AroA - ST5E3PS pro	(428)	-								
<i>S. typhimurium</i> aroA - STYAROA PM pro	(428)	-								
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(428)	-								
<i>Shigella dysenteriae</i> AroA - SDU82268 pro	(428)	-								
<i>E. coli</i> AroA - ECAROA pro	(428)	-								
<i>Shigella sonnei</i> AroA - AF101225 pro	(428)	-								
Consensus	(1486)									

EPSP Synthase CDS protein alignment

		1541	1550	1560	1570	1580	1595	Section 29
Oryza sativa AroA gene - AP002542pro	(392)	-	-	-	-	-	-	
Lolium rigidum AroA - AF349754 pro	(348)	-	-	-	-	-	-	
Z. mays AroA - ZMEPSPSpro	(445)	-	-	-	-	-	-	
N. tabacum AroA partial- M61905 pro	(339)	-	-	-	-	-	-	
Petunia hybrida AroA - PETAROA pro	(517)	-	-	-	-	-	-	
Tomato AroA - TOMAROA pro	(521)	-	-	-	-	-	-	
Arabidopsis thaliana AroA cDNA - AF360224 pro	(522)	-	-	-	-	-	-	
Arabidopsis thaliana AroA gene AIEPSP S	(521)	-	-	-	-	-	-	
B. napus AroA - X51475 pro	(517)	-	-	-	-	-	-	
Agrobacterium CP4 partial AroA sequence	(45)	-	-	-	-	-	-	
Brucella melitensis biovar Abortus AroA - AF326475 pro	(481)	-	-	-	-	-	-	
D.nodosus (VCS1001) aroA - DNEPSS3PS pro	(444)	-	-	-	-	-	-	
Corynebacterium glutamicum AroA - AF114233 pro	(499)	-	-	-	-	-	-	
Pyrococcus abyssi AroA - CNSPAX02 pro	(411)	-	-	-	-	-	-	
S. cerevisiae AroA - Z48179.1 pro	(1500)	PADKPLDDELLSKLERFLVKGAHAAFVPTLEAAAYKPSVT	PVMTISQDKYQWHVV					
S. pombe AroA - AL157734 pro	(1490)	PADIEELPEKKVASVIKALLAN	-KADGGVFLDMAYKPLHTPLMAVASD-LEWKCC					
B. pertussis AroA - BPEAROA pro	(443)	-	-	-	-	-	-	
Aeromonas salmonicida AroA - A18838 pro	(428)	-	-	-	-	-	-	
Haemophilus influenzae AroA - HEAAROAUR pro	(433)	-	-	-	-	-	-	
Haemophilus somnius AroA - HEA3P1C pro	(433)	-	-	-	-	-	-	
Pasteurella haemolytica NADC-D60 AroA - PHU03068 pro	(435)	-	-	-	-	-	-	
Pasteurella haemolytica serotype 2 aroA - PHU89948 pro	(433)	-	-	-	-	-	-	
P. multocida aroA - PMAROA pro	(442)	-	-	-	-	-	-	
Vibrio cholerae AroA - pro	(427)	-	-	-	-	-	-	
Y. enterocolitica AroA - YEP SERC ARO pro	(428)	-	-	-	-	-	-	
Yersinia pestis AroA - YEPAROA pro	(425)	-	-	-	-	-	-	
Klebsiella pneumoniae aroA - KPAROA pro	(428)	-	-	-	-	-	-	
S. typhi AroA - ST5E3PS pro	(428)	-	-	-	-	-	-	
S. typhimurium aroA - STYAROA PM pro	(428)	-	-	-	-	-	-	
Salmonella gallinarum AroA - STYSERARO pro	(428)	-	-	-	-	-	-	
Shigella dysenteriae AroA - SDU82268 pro	(428)	-	-	-	-	-	-	
E. coli AroA - ECAROA pro	(428)	-	-	-	-	-	-	
Shigella sonnei AroA - AF101225 pro	(428)	-	-	-	-	-	-	
Consensus (1541)								

WO 03/013226

PCT/US02/21837

EPSP Synthase CDS protein alignment

	Section 30			
	(1596) 1596	1610	1620	1630
<i>Oryza sativa</i> AroA gene - AP002542 pro	(392) -			
<i>Lolium rigidum</i> AroA - AF349754 pro	(348) -			
<i>Z. mays</i> AroA - ZMEPSPSPpro	(445) -			
<i>N. tabacum</i> AroA partial- M61905 pro	(339) -			
<i>Petunia hybrida</i> AroA - PETAROA pro	(517) -			
<i>Tomato</i> AroA - TOMAROA pro	(521) -			
<i>Arabidopsis thaliana</i> AroA cDNA - AF360224 pro	(522) -			
<i>Arabidopsis thaliana</i> AroA gene AIEPSPS	(521) -			
<i>B. napus</i> AroA - XS1475 pro	(517) -			
<i>Agrobacterium</i> CP4 partial AroA sequence	(45) -			
<i>Brucella melitensis</i> biovar <i>Abortus</i> AroA - AF328475 pro	(481) -			
<i>D. nodosus</i> (NCS1001) aroA - DNEPSP3PS pro	(444) -			
<i>Corynebacterium glutamicum</i> AroA - AF114233 pro	(499) -			
<i>Pyrococcus abyssi</i> AroA - CNSPAX02 pro	(411) -			
<i>S. cerevisiae</i> AroA - Z48179.1 pro (1555) PGSQLMLVHQGVAQFERWTGFKGPFKAIFDAVTKE-				
<i>S. pombe</i> AroA - AL157734 pro (1541) NGLEALVLRQGLASFHLLWTGMTAPFDAYQQKVIE--				
<i>B. pertussis</i> AroA - BPEAROA pro	(443) -			
<i>Aeromonas salmonicida</i> AroA - A16838 pro	(428) -			
<i>Haemophilus influenzae</i> AroA - HEAAROAUR pro	(433) -			
<i>Haemophilus somni</i> AroA - HEA3PIC pro	(433) -			
<i>Pasteurella haemolytica</i> NADC-D60 AroA - PHU03068 pro	(435) -			
<i>Pasteurella haemolytica</i> serotype 2 aroA - PHU89948 pro	(433) -			
<i>P. multocida</i> aroA - PMAROA pro	(442) -			
<i>Vibrio cholerae</i> AroA pro	(427) -			
<i>Y. enterocolitica</i> AroA - YEPSEERCAROpro	(428) -			
<i>Yersinia pestis</i> AroA - YEPAROApro	(425) -			
<i>Klebsiella pneumoniae</i> aroA - KPAROA pro	(428) -			
<i>S. typhi</i> AroA - ST5EAPS pro	(428) -			
<i>S. typhimurium</i> aroA - STYAROAPM pro	(428) -			
<i>Salmonella gallinarum</i> AroA - STYSERARO pro	(428) -			
<i>Shigella dysenteriae</i> AroA - SDU82268 pro	(428) -			
<i>E. coli</i> AroA - ECAROA pro	(428) -			
<i>Shigella sonnei</i> AroA - AF101225 pro	(428) -			
Consensus	(1596)			

FIGURE 3
Page 30